

378
B576

10110

LIBRARY
College of St. Francis
JOLIET, ILL.

University of St. Francis

GEN 378 B576

Bixler

Check Lists for Educational Re



3 0301 00018345 5

M

STANDARD RESEARCH PROCEDURES

EDITED BY
WILLIAM A. McCALL, Ph.D.

CHECK LISTS
FOR EDUCATIONAL RESEARCH
BY HAROLD H. BIXLER

OTHER VOLUMES IN PREPARATION

CHECK LISTS
FOR
EDUCATIONAL RESEARCH

BY
HAROLD HENCH BIXLER, PH.D.

DIRECTOR OF RESEARCH AND GUIDANCE
ATLANTA PUBLIC SCHOOLS
SOMETIME ASSOCIATE IN ELEMENTARY EDUCATION
TEACHERS COLLEGE, COLUMBIA UNIVERSITY

BUREAU OF PUBLICATIONS
Teachers College, Columbia University
NEW YORK CITY

1928

COPYRIGHT, 1928
BY
TEACHERS COLLEGE
COLUMBIA UNIVERSITY

EDITOR'S INTRODUCTION

A FEW centuries ago there was violent opposition to a thoroughgoing application of the scientific method in answering questions within the so-called physical sciences. The scientific method now faces a like extreme opposition in predominantly social realms. Trial-and-error tampering with problems arising in connection with morals, marriage, political institutions, and the like is common enough, but few, if any, of these attempts can qualify as social experimentation. They are usually undertaken passionately and with conviction already established. Predominant power undertakes by force to discontinue each experiment without waiting for its conclusions; and the experimenters, who are seldom notable for meekness, propagandize their theories before they have been established. We thoroughly deserve what our descendants will be saying about us in this connection a few centuries hence.

With respect to the freedom given to dispassionate research, education is fortunate. Such opposition as exists arises more from a doubt as to the feasibility of applying the scientific method to important problems of education than from any intrinsic antagonism. What will happen when the method is extensively applied to the problem of educating the purposes of children, involving, as this inevitably will, the question of character building, religious instruction, and economic and political indoctrination, remains to be seen. One thing is certain. The amount of support given both to the established lines of research and to

these vitally important peripheral phases will depend in large measure upon the adequacy of the research techniques employed by investigators and upon the soundness of their interpretation of resulting data.

The series entitled *Standard Research Procedures* has been established to help secure such adequacy not only for these peripheral problems but also for the multitude of problems still remaining in fields which research has already sampled. It is hoped that the series will encourage the codification, refinement, and origination of research methods generally applicable to research of any kind in any field and particularly applicable to education.

The editor envies Dr. Bixler the authorship of this first volume in the series, not because it is the first but because it treats more types of research techniques than any book previously published, and treats them so usefully that it is destined to become one of those indispensable books kept within reach for constant consultation.

WILLIAM A. MCCALL

*Teachers College
Columbia University
New York City*

ACKNOWLEDGMENTS

THE AUTHOR is grateful to Professor William A. McCall for the suggestion which led to the writing of this monograph, and for inspiration, guidance, and editorial criticism. Professor Arthur I. Gates made helpful suggestions at the inception of the study. Professors Rudolf Pintner, J. Ralph McGaughy, and Percival M. Symonds gave assistance and helpful criticisms during the progress of the study and read the final draft. Many valuable suggestions were received from others, among whom may be named the faculty members who lectured in the Methods of Research course during the year 1926-27, the members of the seminar in Educational Psychology, and many graduate students of Teachers College, Columbia University. Thanks are also due to the students in the Methods of Research course whose judgments are reported in one chapter. It is difficult to express adequately the author's debt to numerous writers of books and articles, from whom ideas have been freely drawn. Among these should be mentioned Carter Alexander, Walter V. Bingham, W. W. Charters, Max Engelhart, Fred M. Fling, Max Freyd, Henry E. Garrett, Truman L. Kelley, Walter S. Monroe, William A. McCall, Giles Ruch, William Schluter, George Stoddard, Percival Symonds, and the Joint Committee of the American Society of Mechanical Engineers. Finally, the author is deeply indebted to his wife, Margaret May Bixler, for assistance at all stages of the study.

H. H. B.

CONTENTS

EDITOR'S INTRODUCTION	v
CHAPTER	
I. CONSTRUCTION OF THE CHECK LISTS AND THEIR USE IN RESEARCH	1
Construction of the Check Lists	2
Application of the Check Lists	3
Use of the Check Lists	4
II. SELECTION AND FORMULATION OF A RESEARCH PROBLEM	5
Selection of the Problem	5
Preparation of Bibliography and Survey of Literature	6
Formulation of the Problem and Outline of Procedure	9
Examination of the Proposed Sources of Data...	11
III. METHODS OF RESEARCH	13
Historical Methods	13
Experimental Methods	15
The Subjects	16
The Materials of Experimentation	18
The Place of the Experiment	18
The Time of the Experiment	19
Equating Groups	19
Control and Measurement of Experimental and Irrelevant Factors	19
Case Method	21
Case-Group Method	22
Investigational Methods	23
IV. TECHNIQUES USED IN RESEARCH	25
Measurement	25
Validity of a Test	26

Reliability of a Test	28
Scaling a Test	30
Test Norms	31
Manual of Directions for a Test	32
Scoring Formulas and Keys	33
Duplicate Forms of a Test	33
Diagnostic Value of a Test	34
Mechanical Features of a Test	34
Administration of a Test	35
Scoring Tests	37
Interviews	38
Questionnaires	40
Selection and Arrangement of Items in a Questionnaire	41
The Form of the Questionnaire	43
Mailing and Follow-up of Questionnaires	44
Judgments or Ratings	45
The Judges or Raters	48
Other Techniques of Collecting Data	50
Checking the Data	50
V. STATISTICAL METHODS USED IN RESEARCH	52
Distributions	53
Measures of Central Tendency	53
Measures of Variability	54
Correlation	55
Reliability	58
Reliability of Tests	60
Miscellaneous	61
VI. THE PRESENTATION OF THE RESEARCH	63
Introduction	63
Statement and Definition of the Problem	64
Presentation of Data	65
Tables	67
Graphs and Charts	69
Citations	71
References	72

CHAPTER	PAGE
Conclusion	73
The Bibliography	75
The Appendix	76
The Form of Presentation	76
English Structure and Form	79
VII. THE APPLICATION OF THE CHECK LISTS IN THE ANALYSIS OF PUBLISHED RESEARCHES	80
The Selection of the Researches	81
Analysis of the Dissertations	83
The Statement of the Problem in One Hundred Dissertations	84
Methods of Research in One Hundred Disserta- tions	85
Techniques of Collecting Data in One Hundred Dissertations	86
Statistical Computations in One Hundred Dis- sertations	89
Methods of Presentation in One Hundred Dis- sertations	93
Judgments of Graduate Students	94
VIII. CONCLUSION	96
BIBLIOGRAPHY	99
APPENDIX: ADDITIONAL ITEMS, NOTES, AND REFERENCES...	107
INDEX	113

TABLES

NUMBER	PAGE
I. Distribution, by Universities, of Dissertations Selected for Analysis	82
II. Analysis of the Statement of the Problem in One Hundred Dissertations	83
III. Analysis of the Methods of Research Used in One Hundred Dissertations	85
IV. Analysis of the Techniques of Collecting Data in One Hundred Dissertations	87
V. Analysis of the Statistical Methods Employed in One Hundred Dissertations	90
VI. Analysis of Methods of Presentation Used in One Hundred Dissertations	93
VII. Judgments of Ninety-eight Graduate Students as to Topics That Should Be Included in a Methods of Research Course	95

CHECK LISTS
FOR
EDUCATIONAL RESEARCH

CHECK LISTS FOR EDUCATIONAL RESEARCH

CHAPTER I

CONSTRUCTION OF THE CHECK LISTS AND THEIR USE IN RESEARCH

THE field of education offers many opportunities for research. Indeed, as educators strive to find a scientific basis for their procedures more and more research is necessary. And yet the number of competent research workers is few, in proportion to the number of problems to be attacked.

Furthermore, much of the research can be successfully prosecuted only in typical field situations—the classroom, the administrative office, the school environment. This fact restricts the available number of research workers, since many of them are engaged in university work. The schools which are accessible to them are often neither typical nor adequate.

It is the purpose of this monograph to present an instrument that will make it possible for more school men and women to participate in scientific research. The method employed was to analyze the procedure that must be followed in conducting a research. The resulting list of steps may serve as a guide to a new investigator, showing him the consecutive steps that he ought to follow, or showing him the alternatives that are open to him.

2 CHECK LISTS FOR EDUCATIONAL RESEARCH

CONSTRUCTION OF THE CHECK LISTS

The first step in this analysis of research was the assembling of items. These items of course constitute the steps or processes of research. For this purpose the following sources were tapped:

1. Standard works on the methodology of research, such as McCall, *How to Experiment in Education* (46),¹ Kelley, *Statistical Method* (40), Pearson, *The Grammar of Science* (58), and others.
2. Dissertations and research monographs. Those which proved fruitful will be found in the Bibliography. Many, of course, yielded no new items.
3. The files of magazines publishing educational, including psychological, researches. The first magazines selected were those reported by the President of the American Educational Research Association, in his annual presidential address (90). He named the following magazines as containing the largest number of articles on research during the years 1924 and 1925:

Journal of Educational Psychology.

Journal of Educational Research.

Elementary School Journal.

School and Society.

School Review.

Educational Administration and Supervision.

Teachers' College Record.

These magazines were then read systematically beginning with the year 1926 and working backward. In general, this method was followed until two successive years yielded no items. The *Journal of Educational Psychology*, the *Journal of*

¹ Numbers in parenthesis refer to the Bibliography, pages 99 to 106.

Educational Research, and the *Teachers College Record* proved to be the most promising sources of techniques, and were read as far back as the year 1920.

A second group of periodicals included the following:

Pedagogical Seminary.

Journal of Experimental Psychology.

Journal of the American Statistical Association.

Psychological Abstracts.

Psychological Clinic.

4. Items proposed by faculty members and graduate students at Teachers College, Columbia University.

As procedures or techniques were discovered they were recorded on 3" x 5" cards. These cards were then assorted, and the Check Lists, as found on the succeeding pages, were constructed.

APPLICATION OF THE CHECK LISTS

The final step in this study was the application of the Check Lists to a number of doctoral dissertations. It should be pointed out, however, that the Check Lists are intended as guides during the process of research, rather than as an instrument to be applied at the completion of a research.

The purpose of this application was to discover the relative frequency with which the various techniques were used. This is significant as an indication of instruments of research needed by future research workers, and candidates for the doctorate in particular. The tables found in Chapter VII will indicate some of the methods and techniques that will be

4 CHECK LISTS FOR EDUCATIONAL RESEARCH

needed. The writer does not pretend to set up a curriculum for research workers. The content of a course in Methods of Research must be planned in the light of courses in Statistical Methods, Measurement, and similar subjects, which are being offered.

One hundred dissertations, which included practically all those published during the years 1924 and 1925, were read. The results have been tabulated and analyzed in Chapter VII. Judgments of graduate students were also obtained to supplement these findings.

USE OF THE CHECK LISTS

The Check Lists have several possible uses, among which we list the following:

Teachers, principals, supervisors, and other educational workers, particularly those who have had little training or experience in research, can use the Check Lists as a guide in carrying through a research. The Lists will no doubt remind investigators of procedures which otherwise might easily be overlooked.

Persons who are directing the researches of others will find the Check Lists a convenient means of checking the progress of investigations. Thus they can guide the work of less experienced investigators at a saving of time and effort to all.

The Check Lists may also prove helpful in constructing a curriculum for a course in Methods of Research, as indicated above.

In conclusion, one or two limitations should be noted. The Check Lists will probably prove more useful to workers in quantitative research than to workers in philosophical research. Furthermore, the special techniques in such fields as administrative research have not been included.

CHAPTER II

SELECTION AND FORMULATION OF A RESEARCH PROBLEM

SELECTION OF THE PROBLEM

1. The research worker must first select the general field in which he desires to work.
2. He should be reasonably well oriented in the field.
3. He should be interested in the field and in the type of problems which are found there.
4. The field should be one which offers promise for research.
5. The research worker should make a hasty inventory of the existing scientific knowledge in the field, by consulting existing summaries (e.g., 75) and by reading available research studies.
6. He should examine this body of knowledge critically for points needing verification or reworking.
7. He should search for gaps in the verified knowledge.
8. He should look for possibilities for the further extension of knowledge.
9. He should read articles which suggest problems needing to be solved, such as those which appear in the *Phi Delta Kappan* (63), *Educational Administration and Supervision* (2), and the *Elementary School Journal* (33).
10. He should seek the suggestions of experts in the field.

6 CHECK LISTS FOR EDUCATIONAL RESEARCH

He should select one or more problems for further consideration, applying to each one the following questions:

11. Are there other problems whose solution must precede the solution of this one?
12. Is the problem interesting?
13. Do persons of differing points of view and differing backgrounds think that the problem is worthwhile?
14. Why is the problem worthwhile?
15. Is the problem sufficiently worthwhile to justify the time and effort that will need to be employed in the study (94)?
16. Do competent persons believe that the investigator possesses a background of training and experience that qualifies him to undertake the problem?
17. Does the investigator possess, or can he obtain, the resources necessary for the research, such as money, time, authority, materials, etc.?

PREPARATION OF BIBLIOGRAPHY AND SURVEY OF LITERATURE

18. Before a problem is chosen and definitely formulated, it is essential to survey the existing literature (46; 70).
19. The research worker should seek for any existing bibliographies, tapping such sources as—
 - a. Studies or investigations of the problem.
 - b. Bibliographies prepared by the United States Bureau of Education, Washington, D. C., such as the Library Leaflet Series, the bibliographies of research in secondary education (96; 97), and others in typewritten form. (For list of printed bibliographies, see the latest number

of *Publications Available*, published by the Bureau.)

c. Lists of bibliographies (1; 52).

d. Mimeographed and typewritten bibliographies prepared by experts in the field in question.

20. Existing bibliographies must be brought up to date.

21. In order to insure a complete bibliography, the investigator should prepare a list of clues, such as words and phrases, to be used as a guide in searching card catalogues, magazines indexes, and other bibliographical sources (1).

22. In building a bibliography, the references should be recorded in correct form, with full details.¹

23. Among the sources that may be consulted in preparing a bibliography are the following:

Library card catalogues.

United States Catalogue and Cumulative Book Index.

Book Review Digest.

Industrial Arts Index.

Poole's Index to Periodical Literature (1802-1906).

Readers' Guide to Periodical Literature.

Readers' Guide to Periodical Literature: Supplement (1907-19).

¹ The research worker should ascertain the form used by the university or publishing house for which the research is being prepared. Each reference should be recorded on a card. See, for example, the following cards prepared by N. L. Engelhardt and published by the Bureau of Publications, Teachers College, Columbia University, New York: Bibliography Card for a Book Reference (Form B). Bibliography Card for a Magazine Reference (Form M). Bibliography Card for Reference to Year Books, Bulletins, etc. (Form Y). Bibliography Card for Reference to School Reports or Surveys (Form S). Complete data should be obtained at the time the card is prepared. Such details as the date of publication and the main office of the publisher should be recorded accurately and uniformly (53, p. 72).

8 CHECK LISTS FOR EDUCATIONAL RESEARCH

International Index to Periodicals (1920 to date).

Annual Magazine Subject-Index (1908 to date).

Annual subject indexes of special (including foreign) magazines in the field, particularly those not indexed in the above sources.

Record of Current Educational Publications (United States Bureau of Education).

Current periodicals.

Journals that publish reviews of articles and books, such as *Psychological Bulletin*, *Teachers Journal and Abstract*, *Journal of Educational Research*.

Public Affairs Information Service.

United States Library of Congress *List of American Doctoral Dissertations*.

Bulletins listing the titles of masters' and doctors' theses, such as the *Register of Teachers College Doctors of Philosophy 1899-1925* and supplements (61); also mimeographed bulletins issued by the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

24. After a reasonably complete bibliography has been assembled, the investigator should select the most important references for preliminary study.
25. After such preliminary surveys, the research worker will be able to select the particular problem for research.
26. He should then read the existing literature in the field more carefully, keeping in mind the following objectives (46, pp. 11f.; 70):

- a.* To obtain a background of knowledge in the field.
- b.* To discover to what extent the problem has already been investigated.
- c.* To obtain material on which to base a summary of existing investigations.
- d.* To locate additional, incidental, or supplementary problems.
- e.* To obtain suggestions as to procedure.
- f.* To note methods of research and techniques that other investigators have found unsatisfactory or unfruitful.
- g.* To obtain suggestions as to sources of material or data.

FORMULATION OF THE PROBLEM AND OUTLINE OF PROCEDURE

- 27. The research worker should make a general statement of the problem, in the form of one or more questions, or declarative statements (53, Chap. 2). Even at this stage the statement of the problem must be regarded as tentative. Changes are desirable, so long as progress is being made (1).
- 28. He should formulate the objectives: primary and secondary (53, Chap. 2).
- 29. He should analyze the major problem into its elements (24; 53, p. 20).
- 30. He should determine the data or evidence required for each of the elements of the problem (24; 53, Chap. 2).
- 31. He should determine the methods of research applicable to each element of the problem and should state the limitations of these methods. Among the possible methods of approach are the following:

10 CHECK LISTS FOR EDUCATIONAL RESEARCH

Historical method.

Experimental method.

Case method.

Case-group method.

Investigation and compilation, using such techniques as

Tests.

Interviews.

Questionnaires.

Judgments or ratings.

32. He should outline the scope of the study (53, p. 21).
33. He should present sufficient information to orient the problem, including summaries of previous or related studies, if necessary (53, Chap. 2).
34. He should, when necessary, state the point of view or philosophy of education on which the study is based, recognizing any assumptions (53, Chap. 2).
35. He should define the terms used or to be used in the research.

When the outline or "blue print" of the problem is complete, the research worker may check it, using the following questions:

36. Has the problem been sufficiently limited, in view of the time and facilities available, to permit of maximum returns (53, p. 21) ?
37. Are the most direct and convenient methods used?
38. Have the proposed methods been carefully outlined in detail?
39. Has the investigator endeavored to think through every step of the proposed research?
40. Can a miniature research or preliminary trial be

carried out to test the proposed methods (94, p. 253) ?

41. Are related problems which are not included clearly stated so as to avoid any possible confusion (53, p. 21) ?
42. Does part or all of the proposed research rest on assumptions? If so, can these assumptions be justified (23, p. 82) ?

EXAMINATION OF THE PROPOSED SOURCES OF DATA

43. The research worker should ascertain if data can be obtained from primary sources.²
44. He should ascertain if the secondary sources are as reliable as the available primary sources (70).
45. He should ascertain whether the greater reliability obtained by tapping the primary sources compensates for the time and effort required (70).

If secondary sources are to be used, the research worker must ask the following questions:

46. By whom were the original data collected (70) ?
47. Under what conditions were the original data collected (70) ?
48. What changes have the data undergone since they were collected (70) ?
49. Are the secondary sources based directly on the primary sources (70) ?
50. Can the data be checked by other secondary sources (70) ?

² Schluter, in *How to Do Research Work*, says, "Primary or original sources are usually more reliable, because the research worker, through personal contact, is aware of their significance in the light of conditions under which they are obtained. He is also aware of limitations and of the inherent nature of the data" (70).

12 CHECK LISTS FOR EDUCATIONAL RESEARCH

Having located the available sources of data, the research worker may then apply the following checks:

51. Are the data available in sufficient quantity to warrant the study?
52. Have the importance and value of the data been judged by the use of criteria set up by the worker?
53. What evidence is there that the data are suitable for the purpose?
54. Are the data capable of unambiguous interpretation?
55. What evidence is there that the data will be sufficiently comprehensive?
56. What evidence is there that the data are sufficiently reliable?
57. Are the proposed data as objective as the field of study permits (53, p. 27) ?
58. Are the data representative, as judged by
 - Parts of the country from which they are to be obtained.
 - Age groups.
 - Sex.
 - Economic groups.
 - Nationalities.
 - Other pertinent considerations.
59. Are all types of available data considered?
60. Have data of little importance or doubtful value been dropped from consideration?

CHAPTER III

METHODS OF RESEARCH

THE worker is now ready to plunge into the research proper. He should make a thorough review of his outline, and then select the particular portions of the following Check List which are pertinent to the various phases of the research.

HISTORICAL METHODS

Since we are concerned primarily with quantitative educational research, this section will not attempt to exhaust the field of historical research.

1. The research worker must determine, after consideration of his problem, whether historical methods must be employed.
2. If the research has to do with an institution or an organization, the investigator must decide whether a knowledge of the history of the institution is pertinent to an understanding of the present situation.
3. If the research is a study of techniques or procedures, he must determine whether a historical background will contribute to the study.
4. In general, he must decide how much historical material it is necessary to present in order to orient the reader and give him a clear understanding of the problem (53, pp. 21f.).
5. The scope of the investigation should be limited

14 CHECK LISTS FOR EDUCATIONAL RESEARCH

so that the work may be done thoroughly (28, p. 36).

6. A topic can be successfully investigated only if the investigator possesses a good working background (28, p. 35).

7. The value of a historical narrative depends upon
a. The quantity and quality of the sources (28, p. 26).

b. The critical skill and constructive ability of the research worker (28, p. 26).

8. All the sources on a given topic must be located and all must be critically evaluated on the basis of the following criteria:

Genuineness (28, Chap. 3).

Authorship (28, Chap. 4).

Time and place of writing (28, Chap. 4).

Independence (28, Chap 5).

9. Whenever possible the original sources should be obtained. A typical selection of original sources for American educational history follows:¹

Laws, ordinances and resolutions: federal, state, local.

Reports of the United States Commissioner of Education.

Reports of state superintendents.

Reports of state boards of education.

Reports of city and county boards of education.

Reports of city and county superintendents.

Minutes and other records of boards of education.

Courses of study: city, county, and state.

¹The author is indebted to Dr. Edward H. Reisner, Teachers College, Columbia University, for some of the items in this section.

Magazines and newspapers.

Textbooks.

Books on pedagogy.

Proceedings of educational associations.

Publications of national and state educational associations.

Publications of the United States Bureau of Education.

Catalogues and other publications of educational institutions.

Barnard's *American Journal of Education*.

Court decisions and records.

Reports of federal and state census.

Commercial and Financial Chronicle.

10. The investigator must be able to read original sources, as well as secondary works, in the language in which they were written (28, p. 28).
11. The investigator should check each original source by other original sources (28, Chap. 3), two independent witnesses being required for certainty (28, p. 103).
12. Notes should be made with care, and should include complete, literal quotations, in the case of important points (28, p. 111).
13. In writing a piece of historical research, the evidence must be cited in order to establish the probability or certainty of statements (28, pp. 108, 171).

EXPERIMENTAL METHODS ²

14. In any experimental study the total net change in the trait or traits in question produced by irrelevant factors must be negligible, or the amount of

² The best treatment of experimental methods is found in McCall, *How to Experiment in Education* (46, pp. 14-36). The following statements are

16 CHECK LISTS FOR EDUCATIONAL RESEARCH

- such change must be measured and discounted by the application of a control experimental factor (46, p. 21).
15. The one-group experimental method is valid where the change produced by an experimental factor is not conditioned significantly by any preceding factor, and where the change effected by each experimental factor is measurable in equal units (46, p. 21; 53, p. 30).
 16. The equivalent-groups method is valid where it is possible to equate groups (46, p. 29; 53, p. 30).
 17. The rotation method is valid where the change produced by an experimental factor is not conditioned significantly by any preceding experimental factor (46, p. 31; 53, p. 32).
 18. The experimenter should make plans for keeping an experimental log or record of events, with dates, and should include irrelevant factors and unusual conditions (46, p. 79; 23, p. 54).

THE SUBJECTS

19. The experimenter should have a carefully thought out plan for selecting the subjects, that is, the group to which the experimental factors are to be applied (46, pp. 37 ff.).

quoted or paraphrased from his discussion. Experimental methods may be grouped under three principal types, which may be described as follows:

The one-group experiment is that in which an experimental factor or factors has or have been applied to, or subtracted from, one group, thing, or individual.

The equivalent-groups experiment is that in which two equivalent groups are selected, to one of which is applied an experimental factor, and to the other another experimental factor, or else no factor at all.

The rotation experiment is that in which the experimental factors are applied successively, but in reverse order, to the groups. Thus, in the case of one group, there is applied one factor, then the second; in the case of the second group, the second factor is applied first.

20. If it is proposed to use statistical formulas whose validity depends upon the assumptions of random sampling, the experimenter must take precautions to assure that these assumptions are met (100, Chap. 13; 50, Chap. 16).
21. When the composition of the population to be studied is known, the sample must be so selected as to include all the essential elements (19, pp. 112 ff.).
22. When the composition of the population is unknown, the experimenter should select at random sample groups and continue selecting groups until comparison of the samples taken reveals a definite tendency for variations in one direction to occur as frequently as variations in another direction (19, pp. 118 ff.).
23. The method of sampling should be described in full so that the research may later be evaluated.

The experimenter, in selecting subjects for the experiment, must check on the following points (46, Chap. 3) :

24. The subjects must be appropriate to the experimental methods.
25. The subjects must be appropriate to the tests to be used.
26. The subjects must be appropriate to the experimental factors.
27. The subjects should be representative as to age, sex, grade, intelligence, etc.
28. The experimenter should select subjects who may reasonably be expected to be available throughout the period of the experiment.
29. The number of subjects must be sufficiently large

18 CHECK LISTS FOR EDUCATIONAL RESEARCH

to assure reliability of the results. (See section on Reliability, Chap. V.) Sometimes a preliminary trial experiment is needed to determine the approximate number.

30. The number of subjects must be sufficiently large to allow for losses or eliminations on account of absences from one or more tests, inability to pair, and other reasons. If there is a control group it is usually desirable to select a much larger number of subjects for it than will be finally needed, so that all members of the experimental group may be utilized in pairing.

THE MATERIALS OF EXPERIMENTATION

31. The materials must be appropriate to the subjects.
32. The materials must be appropriate to the experimental method.
33. The cost of the materials must be within reasonable limits.
34. The materials must be appropriate to the conditions under which the experiment will be conducted, such as the time available and the qualifications of the persons who are to use the materials.
35. The materials must not present irrelevant or variable factors that might invalidate the experiment.

THE PLACE OF THE EXPERIMENT

36. The place or places in which the experiment is to be conducted must be typical of the situations to which the results of the experiment are expected to apply.
37. This representative character may be planned with reference to parts of a city, parts of a county or state, parts of the United States, etc.

38. The place and its environment must be appropriate to the experimental method and the experimental factors (46, Chap. 4).
39. The place or its environment must not introduce irrelevant or variable factors that might invalidate the experiment (46, Chap. 4).

THE TIME OF THE EXPERIMENT

40. The experimenter must determine whether the time of the day, of the month, or of the year is a significant factor in the experiment (46, p. 74).
41. In deciding upon the time length of an experiment, the principle to be kept in mind is that one should aim to secure the maximum effect of the experimental factors with a minimum effect from irrelevant or variable factors (46, p. 74).

EQUATING GROUPS

When groups are to be equated,³ one or more of the following bases may be used:

42. Random selection, that is, chance arrangement (46, p. 41; 100, Chap. 13).
43. General ability (46, p. 42).
44. Initial status in the experimental trait (46, p. 49).
45. Composite of several test scores (46, p. 51).

³ McCall, in *How to Experiment in Education* (46), says, "Equating is usually on just one basis, first, because every increase in the number of bases employed reduces the number of pupils who can be satisfactorily paired from a given total number of pupils; and, second, because equating on one basis tends to make the groups have approximately equivalent means and variabilities on any other basis, even though particular pupils do not pair on all the bases."

The *time* for equating may well receive careful consideration. McCall says, "Where the actual physical shifting of subjects is not to take place, and, in addition, where the subjects are not to be segregated for purposes of applying the experimental factor, delayed equating is preferable to early equating."

20 CHECK LISTS FOR EDUCATIONAL RESEARCH

46. Rate of growth or change in the experimental trait (46, p. 55).
47. Multiple bases, such as chronological age, mental age, and initial status (46, p. 56; 32).

CONTROL AND MEASUREMENT OF EXPERIMENTAL AND IRRELEVANT FACTORS

48. If tests are needed in the experiment, see section on Measurement, Chap. IV.
49. The experimental conditions or situations must be appropriate both to the problem and to the subjects (46, Chap. 4).
50. The experimenter should have or obtain the authority necessary for the control of experimental conditions (46, Chap. 4).
51. The experimenter should measure the amount of the experimental factors as accurately as possible (46, Chap. 5). See section on Measurement, Chap. IV.
52. The experimenter should measure the amount of change in the trait in question produced by the experimental factors (46, Chap. 5).
53. The amount of change in the experimental subjects due to irrelevant factors should be eliminated, equated, or accurately measured and discounted (46, Chap. 5).

Among the constant irrelevant factors, which may or may not be significant, are the following:

54. Maturing or growth of subjects (46, p. 63).
55. Bias of experimenter or assistants (46, pp. 64f.).
56. Capability of assistants (46, pp. 67f.).
57. Enthusiasm or unwillingness of assistants (46, p. 68).

58. Bias of the subjects (46, pp. 70f.).
59. Differences in time allowance (46, pp. 72f.).
60. Too short or too long a time allowance (46, p. 74).
61. Differences in transfer of knowledge, enthusiasm, and the like, among subjects (46, pp. 75f.).
62. Bias of the tests (46, p. 77).
63. Physical environment, such as temperature, home surroundings, and the like (46, p. 78).
64. Unnatural conditions, or lack of adjustment of the subjects to the situation.
65. The experimenter must keep accurate, detailed, and dated records of the methods of applying the experimental factors, possible irrelevant factors, and other pertinent data (46, p. 79).
66. The research worker should determine whether any other method, such as partial correlation, may be used to supplement or check on the experimental method (46, pp. 224-44).

CASE METHOD

The research worker who plans to use the case method must first determine what procedures he will follow and then select the appropriate techniques. Among the procedures that may be used are the following (35; 42; 59; 64; 71) :

67. He may obtain a history of the case, gathering data as to the developmental record, health, social, educational, and economic history of the person (64) by using the interview technique (see section on Interviews, Chap. IV).
68. He may study the background of the case, including the social, economic, and educational history of the family (64). The techniques used may include the interview (Chap. IV), the questionnaire

22 CHECK LISTS FOR EDUCATIONAL RESEARCH

Chap. IV), and the analysis of records and documents, such as diaries, school records, employment records, letters, and other similar records (42, Chap. 50).

69. He may study the present status of the case, using—
- a.* Measurement of physical, educational, and mental ability or achievement (59; see section on Measurement, Chap. IV).
 - b.* Judgments or ratings (see section on Judgments or Ratings, Chap. IV).
 - c.* Interviews with physician, psychiatrist, employer, teacher, family, friends, and the person himself (64).
 - d.* Observation, including photographs, motion pictures, and the like (42, p. 100; 53, p. 36).
 - e.* Records, including medical examination, school record, court record, and the like (42, Chap. 26; 64).
 - f.* Questionnaires, to be answered by the individual, parent, teacher, employer, or other persons who know the case (see section on Questionnaires, Chap. IV).

CASE-GROUP METHOD

70. Allied to the case method is the case-group method.⁴
71. If case-groups are to be studied, the research worker must determine what qualities or characteristics shall be chosen to differentiate the groups (74, p. 340; 54).

⁴ Snedden (74) says, "This [the case-group method] in essence consists in the isolation for scientific study of groups of human beings fairly homogeneous as respects one or several qualities."

72. He must determine how the information about these qualities can be obtained.
73. He must set up a plan for isolating the groups of persons for study.
74. If two or more case-groups are to be compared, the research worker must collect data, using one or more techniques (Chap. IV).

INVESTIGATIONAL METHODS

In many researches, the method of attack to be used may be termed investigational. The word is used here to include all methods of research that are not classifiable as historical, experimental, case, or case-group methods. Investigations may be classified as descriptive and causal investigations (46, p. 5). The purpose of a descriptive investigation is to describe a situation as accurately, objectively, and quantitatively as possible. The research worker undertaking a descriptive investigation may check on the following points:

75. He must choose the situations he wishes to describe (15).
76. He must ascertain what types of data are available and obtainable (53, Chap. 3).
77. He must decide what types of data are needed in his research.
78. He must decide what techniques he will employ in collecting the data (Chap. IV).
79. He must determine, in a general way, the statistical methods that he will use in treating the data, since they may determine or affect the form in which the data are collected (Chap. V).

The methodology of the causal investigation is the methodology of the equivalent-groups experiment

24 CHECK LISTS FOR EDUCATIONAL RESEARCH

worked backward (46, pp. 211 f.). The essential procedures in a causal investigation are as follows:

80. The investigator formulates hypotheses as to the probable causes of an observed effect (46, p. 213).
81. He collects data, using such techniques as measurement, interviews, questionnaires, judgments, and the like (Chap. IV; 46, pp. 214 ff.).
82. He determines the significance of causes (46, pp. 219 ff.).
83. He studies, statistically, the correlation⁵ and inter-correlation between causes and effects (46, pp. 224 ff.; 13, pp. 11-18).
84. He makes a final evaluation of causes (46, pp. 233 ff.).

⁵ For formulas and illustrations of the use of partial correlation in causal investigations, see Franzen, Raymond, "A Comment on Partial Correlation." *Journal of Educational Psychology*, 19:194-97, March 1928.

CHAPTER IV

TECHNIQUES USED IN RESEARCH

IN THIS chapter there will be presented some of the techniques that may be used in collecting data; namely, measurement, interviews, questionnaires, and judgments or ratings. In collecting data the research worker must anticipate the statistical treatment which he proposes to apply to the data. For example, if it is proposed to use statistical formulas whose validity depends upon the assumptions of random sampling, he must take precautions to assure that these assumptions are met at the time the data are collected (50, Chap. 16; 100, Chap. 13).

MEASUREMENT

The research worker who plans to use measurement techniques will choose the measuring instruments appropriate to the problem. In this chapter, the word "test" is used to refer both to standardized tests, for sale by many publishers, and to the informal tests which the research worker may have to construct himself.

If a certain type of test is needed, the investigator should ascertain what tests are already available.¹ If he finds tests that appear to be suitable, he should evaluate each one, using the Check Lists in the nine sections following. He should beware of the fallacy of

¹ See, for example, *Bibliography of Tests for Use in Schools* (World Book Company).

26 CHECK LISTS FOR EDUCATIONAL RESEARCH

assuming that two tests measure the same thing because they are called by the same name (13, p. 20). The name does not always indicate the content.

If the research worker finds it necessary to construct his own tests, the same sections will prove helpful. The novice in the field of measurement will find it desirable to obtain a background, and especially a technical vocabulary, by referring to the books by such writers as Bingham and Freyd (6), Kelley (41), McCall (46), Monroe (51), Otis (57), Ruch and Stoddard (67), and Symonds (84). The author is indebted to these and other writers for many items. For convenience, the number of references per item has been limited.

The research worker who has chosen his tests should refer immediately to the section on Administration of Tests.

VALIDITY OF A TEST

1. In evaluating a test, the chief criterion is validity, that is, the adequacy and detail with which the test measures the trait, function, or school subject that it purports to measure (84, p. 279).
2. In constructing achievement tests, most of the validation of the test must be accomplished by means of the choice of suitable material for the test (84, p. 280).
3. Concerning the principles underlying the choice of material, there is some difference of opinion, but six methods or criteria may be safely given, as follows (84, p. 280) :

“Criterion of extrinsic use.

“Criterion of errors.

“Composite of textbooks.

"Composite of requirements in course of study.

"Composite of teachers' examinations.

"Judgment of experts." ²

4. After at least twice as many test items are assembled as will be kept permanently, there should be an experimental try-out of these items (67, p. 329).
5. The validity of the individual test items should be determined (47; 67, p. 319).
6. The per cents of pupils passing each item grade by grade or age by age should be computed (67, p. 331).
7. In constructing a test, only a small number of the items which are missed by all pupils should be included in the test (67, p. 45).
8. In constructing a test only a small number of the items which are answered correctly by all pupils should be included (67, p. 45).
9. Test items should show a consistent increase in the percentage of successes with successive age or grade levels (67, pp. 45, 319 ff.).
10. Test items should be free from ambiguities and other irrelevancies (46, p. 89). Ambiguities may be discovered by a study of items frequently missed, especially in recall and multiple choice tests (14).

² Ruch and Stoddard (67, p. 304) present the most frequently used methods for the original selection of items. In addition to those above, the following methods are listed:

"Use of rating scales in setting up criteria.

"Correlations with school marks or other measures of school success.

"Increase in percentage of successes with successive ages or grades.

"Correlations with previously validated measures.

"Differential scores shown by two groups known to be widely separated upon a scale of ability.

"Logical or psychological analysis.

"Correlations with tests of other intellectual, non-intellectual, or educational abilities."

28 CHECK LISTS FOR EDUCATIONAL RESEARCH

11. The effect upon validity of doubling the length of the test may be computed (7, p. 90).
12. The validity of a test as a whole may be determined by obtaining the opinions of competent judges as to the functions measured, and by the correlation of scores on the test with independent criteria (41, pp. 29 f.).

RELIABILITY OF A TEST

13. A test is reliable when it measures accurately, that is, when applications of equivalent forms yield scores that are practically identical (84, p. 287). Other things being equal, the greater the number of test items, the more reliable the test (84, p. 290).
14. Reliability is most frequently reported in terms of reliability coefficients (67, p. 52).
15. In building a test, the test items should be divided so as to obtain at least two equivalent forms, in order that reliability coefficients, that is, coefficients of correlation between the scores of a group of pupils on the two forms, may be obtained (67, p. 338).
16. Objectivity in scoring a test, that is, freedom from judgment or opinion, increases the reliability of the test (84, p. 289).
17. Reliability of a test is increased if the test is evenly scaled, that is, if items are equally spaced in difficulty (84, p. 290).
18. Reliability is increased if mechanical defects, such as catch questions, faulty choice of words, poor sentence structure, and inadequate directions, are eliminated (67, p. 53).
19. Reliability is decreased if the type used in printing

the test is improperly chosen, if the typographical arrangement is poor, or if the paper used is unsuitable.

20. Reliability is decreased by items which show erratic changes in percentage of successes from grade to grade or from age to age (67, p. 53).

21. According to Ruch (65) the following facts should be reported concerning the reliability of a test, and judgment as to reliability should be based on them:

“Coefficient of reliability.

“Standard deviations of scores on which the correlation is based.

“Number of cases involved in the calculation of the correlation.

“Representative character of the sampling.

“Age and grade groups included in the population.

“Order of giving the test forms.

“Probable error of a score.

“Probable error of a true score estimated from an obtained fallible score.”

22. The degree of reliability that should be set up as a goal cannot be arbitrarily determined.³

23. If the reliability coefficient of a test in process of construction is not as high as desired, the Spearman-Brown formula will indicate approximately

³ Ruch and Stoddard (67, p. 56) say that reliability coefficients of 0.80 to 0.89 are “fairly adequate for individual measurements.” Symonds (84, p. 299) urges that “One should aim to get a reliability of 0.90 in all cases where the results of the testing are to be used seriously in the administrative or guidance functions of the school.” Kelley (41, p. 211) concludes that if individual diagnoses are to be made, a reliability coefficient of 0.94, as found from a one-grade range of talent, is needed. For group measurement purposes, a reliability coefficient of 0.50, when determined from a one-grade range of talent, may be sufficiently high, depending upon the examiner’s purpose (41, pp. 210 f.).

30 CHECK LISTS FOR EDUCATIONAL RESEARCH

how much the test must be lengthened to secure the desired reliability (29, p. 269).

SCALING A TEST

24. In constructing a test, the relative merits of different methods of scaling tests should be carefully considered, the specific research at hand being kept in mind (46, pp. 94 ff.; 84, pp. 300 ff.).
25. In most tests the items should be arranged in order of difficulty, beginning with the easiest (84, p. 301).
26. To be a true scale, items should be arranged with equal increments of difficulty; and zero on the scale should represent "just not any" of the thing being measured by the test (87, p. 294). It is permissible to disregard this feature of scaling, provided other general cautions are observed (84, pp. 300 f.).
27. It is probably unnecessary to weight test items on the basis of difficulty (38).
28. There should be some items in the test so easy that no pupil will make a zero score (46, p. 92).
29. There should be some items in the test so difficult that no pupil will make a perfect score (46, p. 92).
30. There should be no undistributed scores at any point of the scale (84, p. 301).
31. In general, a test should be of such difficulty as a whole that the mean score of a class is equivalent to about 50 per cent of the total possible score on the test (84, p. 301).
32. Care should be taken not to omit any step in the process of scaling and standardizing a test (46, pp. 130 ff.; 67, pp. 301-75).

TEST NORMS

33. A test builder should provide data for the interpretation of scores or records by presenting one or more of the following:
- a. Grade norms (67, p. 344).
 - b. Tables for transmuting crude scores or the final total scores obtained on a test into grade (G) scores, that is, scores in terms of grades and tenths of a grade ⁴ (48, Chap. 3; 9, p. 181).
 - c. Age norms (67, p. 345).
 - d. Percentile norms (67, p. 347; 46, p. 95).
 - e. T scores (6, pp. 98 ff.).
 - f. Sigma (or other deviation) position in a defined group (41, p. 12).
34. In the selection of subjects upon whose performance norms are to be based, chronological age, mental age, and sociological status should be considered (18).
35. The test builder should report the number of cases on which norms are based.
36. The test builder should report the character of the population on which the norms are based.
37. In the case of grade norms, the time of year when the tests were given should be reported.
38. Norms based on tests given in mid-year may be adjusted to read as of the beginning or the end of a school grade (48; 9, p. 182).
39. Approximate norms may be obtained by calibrating a new test against an existing standardized test, using the line of relation procedure (4; 57, pp. 104 ff.).

⁴ On tests published by the Public School Publishing Company, such scores are called B scores,

32 CHECK LISTS FOR EDUCATIONAL RESEARCH

MANUAL OF DIRECTIONS FOR A TEST

40. The directions to those who are to administer the test should be as brief as is consistent with adequate understanding.
41. These directions, together with instructions to those who take the test and other information as to purpose, validity, reliability, and norms, should be included in a manual of directions (67, p. 375).
42. The instructions to those who are to take the test should be worded in language intelligible to all who may be tested (46, p. 91).
43. The instructions should be written in the order in which they are to be executed (46, p. 92).
44. The instructions should be arranged in units, so that after a unit of instruction has been given, the persons being tested may have an opportunity to carry out the instructions (46, p. 92).
45. The instructions should include information as to the number of parts in the tests or the number of pages the person or group may work through (84, p. 307).
46. The instructions should include a warning to turn the page, or to refrain from turning the page, as the test may require (84, p. 307).
47. The instructions should state the mechanics of answering the questions in the test; for example, underlining, writing numbers, etc. (84, p. 307).
48. In true-false and multiple-choice tests the preferred instruction appears to be, "Do not guess" (66).
49. A fore-exercise or preliminary sample of the test is desirable (46, p. 91). The instructions and fore-exercises should be tried out by the test

builder on a preliminary group in order to discover ambiguities, and to determine other necessary changes and additions.

50. The instructions should include a statement as to the amount of time available, and, in some instances, a warning not to dwell too long on any one item (84, p. 307).
51. The persons being tested should be told what to do in case they finish before time is called (84, p. 307).

SCORING FORMULAS AND KEYS

52. A test should be accompanied by scoring instructions, preferably in a manual of directions.
53. A comprehensive answer key should accompany the test (93, Chap. 6).
54. In true-false tests, the method of scoring Rights minus Wrongs appears to be best (66; 98).
55. The best method of scoring multiple-choice and best-answer tests appears to be Number Right (38).
56. The arrangement test, that is, a test in which the subject is to arrange items in a certain prescribed order, has a special scoring formula (7, p. 21).
57. The matching test is scored by counting the Number Right (84, p. 36).
58. If a test yields two scores, such as measures of quality and of rate, it is possible, if desired, to combine them into a single score (22; 31).

DUPLICATE FORMS OF A TEST

59. If needed in the research, two or more duplicate forms of a test may be constructed, unless the subject has very limited content (84, pp. 305 f.),

34 CHECK LISTS FOR EDUCATIONAL RESEARCH

60. The various forms should be equivalent, that is,
- a. The test items should be random samples of a larger number of valid items (67, p. 65).
 - b. There should be no duplication of items (67, p. 65).
 - c. The average difficulties of the forms should be equal (67, p. 65).
 - d. The various forms should have, within reasonable limits, the same standard deviations and other measures of variability (67, p. 65).
 - e. The various forms should be equal in difficulty at all points on the scale, so that scores of individual pupils may vary as little as possible from form to form (67, p. 66; 84, p. 306).

DIAGNOSTIC VALUE OF A TEST

61. Whether a test should have diagnostic value or not depends on the purpose of the research (67, p. 64).
62. The test builder must determine whether the test must be sufficiently reliable for individual, as contrasted with class, diagnosis (67, p. 19).
63. For a test to meet the demands of genuine diagnosis in a practical situation, stringent requirements must be observed (67, p. 19).

MECHANICAL FEATURES OF A TEST

64. A test should be so constructed that the subject's reactions may be recorded readily, that is, by a word, a number, or a check (46, p. 90; 93, Chap. 2, 3).
65. The subject's reactions should be controlled as to spatial location on the page (46, p. 90; 93).
66. When the percentage of individuals who make a

certain response to a test item by marking printed answers is used as a basis for conclusions as to absolute amounts of knowledge or preference shown, allowance should be made for the effect of the position of the printed responses, since there is a constant tendency to mark the responses printed above $\left(\text{for example, } \frac{\text{more}}{\text{less}} \right)$ and those printed to the left (yes—no) (44).

67. In constructing test items, careful attention should be given to the composition of statements; for example, in true-false items one should avoid “specific determiners” such as “always” and “never” (93, Chap. 7, 8, 9; 43).
68. The time limit should be determined experimentally (67, p. 339).

ADMINISTRATION OF A TEST

69. A test should be administered by trained persons, known to be efficient and accurate.
70. Whenever possible, all tests should be administered by the same person.
71. If several tests are being compared, care should be taken to rotate them, so that no one of them may have any possible advantage (46, pp. 31ff.).
72. The examiner should be familiar with the manual of directions and with the content of the test.
73. Every effort should be made to obtain good working conditions with respect to light, temperature, and environment (48, Chap. 2).
74. The examiner should anticipate and take steps to eliminate any possible distractions (48, Chap. 2).
75. Before beginning to administer a test, the examiner should assure himself that all necessary

36 CHECK LISTS FOR EDUCATIONAL RESEARCH

materials, such as a watch with second hand, pencils, erasers, the manual of directions, and the correct number of test blanks, are at hand (48, Chap. 2).

76. The examiner must secure and hold the attention of the individual or group to be tested.
77. The examiner should strive to obtain a good mental attitude on the part of the individual or group to be tested, and should avoid nervousness or tension (23, pp. 124-39).
78. Care should be taken to avoid any motivation for the test not provided in the directions (48, Chap. 2; 23, pp. 124-39).
79. Before the test materials are distributed, the individual or group should be instructed what to do on receiving test blanks.
80. The examiner or an assistant should see that all members of the group have pencils, test blanks, and all other necessary material.
81. The examiner should allow the individual or group being tested sufficient time to record the desired identifying information on the test blanks.
82. The examiner should observe carefully all the directions provided with the test, particularly those in regard to reading of directions to persons being tested, questions that may be asked, assistance in case of misunderstanding of directions, time keeping, collection of papers and the like, even though they conflict with any general directions above (48, Chap. 2).
83. As soon as the persons taking a test have begun work, the examiner or an assistant should move quietly about the room to see that identification blanks have been filled out correctly, and that pu-

pils are proceeding in accordance with the directions for the test (48, Chap. 2).

84. The examiner should collect all papers, being sure that none are left in the hands of the persons being tested.
85. A record should be made of any interruptions, distractions, or other factors that might influence the results.
86. The examiner should endeavor to determine whether lack of experience in the technique of taking the tests will affect the results, and if so, provide systematic training or practice in advance (84, pp. 295 f.).

SCORING TESTS

87. The scorers should read the special rules governing the scoring of the test before beginning to score the papers (84, p. 81).
88. Consistency in scoring can only be obtained by using answer keys and by following instructions for scoring and for computing derived scores (84, pp. 80 f.).
89. Where scoring is not wholly objective, as in the recall type of test, a careful record should be made of acceptable and unacceptable answers.
90. Scoring can often be expedited by the use of mechanical devices, such as transparent paper or celluloid, perforated sheets of paper or cardboard, or printed strips of paper which can be superimposed on the test blanks (67, p. 58; 93).
91. When a test is composed of several parts or sub-tests, the workers should score the same sub-test in all papers before passing to the next sub-test (84, p. 80).

38 CHECK LISTS FOR EDUCATIONAL RESEARCH

92. The scoring procedure should be systematized; for example, by having a group of scorers score all papers under supervision (84, p. 80).
93. A uniform plan of checking rights, wrongs, and omissions should be followed; for example, rights may be marked (—), wrongs (×) and omissions (0) (48 Chap. 2; 84, pp. 80 f.).
94. In doubtful cases, the scorer should ascertain the intent of the examinee (84, p. 81).
95. All papers should be rescored, preferably by another person.
96. All arithmetical computations, such as subtraction in counting rights minus wrongs, and additions of scores on sub-tests, should be checked, preferably by another person (84, p. 81).

INTERVIEWS

There are many types of research in which the interview technique is employed. Frequently the research worker himself will do the interviewing. Even if the interview is employed as a supplementary technique, the investigator will do well to perfect a technique of interviewing.

97. The questions to be asked during an interview should be carefully prepared in advance, and should be submitted to a group of critics, such as the prospective interviewers, for criticism (20).
98. The research worker should then rewrite the questions in the light of the criticisms offered, and should try them out on a sampling of the group to be interviewed (20; 70, p. 34).
99. He should record and analyze the replies to obtain hints as to further modifications needed to make the meaning clear (20).

100. The questions should then be rewritten in final form, the aims being clarity and maximum drawing power (20).
101. If interviewers have to be employed, the research worker should select persons who are qualified by experience, or he should train them for the work (20).
102. In training an interviewer, the research worker should have him conduct a few interviews alone and write his reports. The research worker should then discuss the interview and report with the interviewer, checking on his report to make sure that he does not add important oral comments or data not included in the report (20).
103. The interview should be prearranged by letter or telephone (70).
104. In arranging a time and place for an interview the convenience of the person to be interviewed should have first consideration, but other elements, such as the suitability of the time and place for the purposes of the research worker must be considered.
105. The interview should be opened in such a way as to interest the person being interviewed in the object of the interview.
106. The interviewer should gain the confidence of the subject or else he should return at a later time.
107. A good interviewer holds his informant to the point (20).
108. A good interviewer is persistent in digging in after information with questions until he has obtained all the information he can get (20).
109. A good interviewer is facile in guiding conversation (20).

40 CHECK LISTS FOR EDUCATIONAL RESEARCH

110. The interviewer should stick to the subject or purpose of the interview (70, p. 35).
111. The interviewer should use tact in securing private or confidential information (70, p. 35).
112. The interviewer should avoid fruitless discussions and arguments over differences of opinion (70, p. 35).
113. Data should not be recorded during the interview, if it embarrasses the person being interviewed (70, p. 35).
114. The method of attack should be flexible, not stereotyped, and it ought to be adapted to the person being interviewed (70, p. 35).
115. The interviewer should not tire the subject by a long interview (64; 70, p. 35).
116. When reactions of various persons are to be compared, the interview should be standardized, to some extent, by the use of precisely the same questions.
117. A good interviewer writes detailed reports (20).
118. A good interviewer reports some of the significant responses in the words of the subject (20).

QUESTIONNAIRES

119. A questionnaire should be used only when absolutely necessary; for example, when the data cannot be obtained in any other way.
120. Before deciding to construct a questionnaire, the research worker should try to ascertain if the persons to be addressed are in a position to know the answers to the questions.
121. The research worker should also try to ascertain whether the persons to be addressed are willing to supply the data requested, keeping in mind

the nature of the data, the prestige of the investigator, and the endorsements which it is possible to obtain.

122. A good questionnaire is of such a length that the recipients may be reasonably expected to give the amount of time necessary for accurate answers.
123. A good questionnaire deals with matters that are worth investigating and that will seem to the recipients to be worth investigating (94).
124. A good questionnaire, although requiring only brief replies, will stimulate supplementary communications (94).

SELECTION AND ARRANGEMENT OF ITEMS IN A QUESTIONNAIRE

125. The meaning of the questions in a questionnaire should be so clear that only one interpretation is possible (70, p. 38).
126. The questions should be so worded that the replies evoked will not be ambiguous.
127. Unusual or technical terms should be so defined that they will convey the same meaning to all persons addressed.
128. The questions should clearly indicate whether facts or opinions are requested.
129. If facts are requested, it should be clear whether the facts are to be exact or estimated.
130. If opinions are requested, the research worker should assure himself of their probable value.
131. The questions should be so worded that they can be answered with a minimum of writing, that is, by checking, underscoring, etc. (70, p. 38; 56).
132. The questions should follow a logical sequence in

42 CHECK LISTS FOR EDUCATIONAL RESEARCH

order to stimulate interest and facilitate answering (70, p. 41).

133. The questions should be concise (6, pp. 145 ff.).
134. The questions should cover all the data needed so that a second questionnaire will not be required.
135. The questions should be stated in such a way that as much handling of the data as possible may be done by the research worker rather than by the person questioned (70, p. 42).
136. Questions should be inserted that will serve as a check on other questions. For example, "Age" may be checked by "Date of birth"; an opinion about a person or thing, by questions on specific traits or points (70).
137. If a vitally important point is to be established, a fairly large number of questions should be framed, and they should be "interlocked" in such a way that conclusions may be based on a consensus of answers rather than on one isolated statement (99).
138. Other things being equal, the investigator should use as few questions as possible, in the light of the nature of the problem, the character of the questions, and the degree of responsiveness reasonably to be expected (70, p. 39).
139. The questions should be so framed that the answers will lend themselves to tabulation and statistical treatment (94).
140. Objective, quantitative information should be the aim in most questionnaires.
141. Questions should be so phrased as to avoid suggesting bias.
142. The questionnaire as a whole should be checked

to make sure that it meets the objectives of the research.

143. After the questionnaire has been prepared in tentative form, the research worker should submit it to a group of critics for suggestions (20).
144. The questionnaire should then be tried out on a sampling of the group to whom it is to be addressed (20).
145. The research worker should record and analyze the replies to this preliminary questionnaire, and make modifications that will clarify the meaning and obtain the maximum drawing power (20).

THE FORM OF THE QUESTIONNAIRE

146. The size of the questionnaire should be convenient for handling and filing (19, pp. 148 ff.).
147. If spacing is not adequate rules should be employed to separate the main categories (19, pp. 148 ff.).
148. Spacing should aid the eye and allow sufficient room for full details (19, pp. 148 ff.).
149. Identifying data should be conspicuously placed at the top of the questionnaire (19, pp. 148ff.).
150. Spaces for answers, especially if the answers are in the form of checks, numbers, etc., should be near the right-hand margin of the page to permit of ready tabulation (6, p. 149).
151. Items should be so spaced that none will be overlooked.
152. Questions should usually be phrased in the third person.
153. The possible typographical forms, such as typewritten, mimeographed, and printed, should be carefully considered (70, p. 44).

44 CHECK LISTS FOR EDUCATIONAL RESEARCH

154. The general appearance and form of the questionnaire should be such as to motivate the person addressed to give complete and accurate answers (6, p. 145; 19, pp. 148 ff.).

MAILING AND FOLLOW-UP OF QUESTIONNAIRES

155. When it is necessary to select the persons to be addressed from among an available number, they should be selected in such a way as to avoid bias.
156. The number and character of the persons addressed should be such that errors of judgment and of statement may be balanced, thus yielding a net total that approximates the truth (99).
157. Mere quantity of replies will not make for quality (94).
158. The number of persons addressed should be large enough to allow for losses, and yet to insure a sufficient number to satisfy statistical criteria of reliability (see section on Reliability, Chap. V). The approximate number of replies required may be forecast by a preliminary trial (94, p. 255).
159. The questionnaire should be accompanied by a letter which states the purpose of the research and makes a courteous appeal to the interest of the addressee, on the basis of the importance of the inquiry, or of his unique and authoritative position (70, p. 37). In most cases, the addressee should be informed that he will receive a copy of the published results if he desires it.
160. A stamped, addressed envelope should be inclosed.
161. Questionnaires should be mailed to reach their destinations at a time when replies may be expected. Busy seasons, vacation periods, and the like should be avoided (70, p. 43; 88).

162. Follow-up letters and telegrams should be used to obtain the return of all possible questionnaires. (Toops (88) found five follow-up messages worthwhile.)
163. A daily graph of the number of questionnaires returned should be kept as a check on the success of follow-up letters, and also as a guide to the proper time to send another follow-up letter (88).
164. In order to determine whether the questionnaires returned constitute a random sample of the total, it is advisable to tabulate, from time to time, the replies on significant items. If little variation appears from week to week there is reasonable assurance that no selective factor is operating (88).

JUDGMENTS OR RATINGS

165. The research worker who contemplates the use of ratings should recognize the lack of validity due to many sources of error, such as
 - a. Halo effect, or the tendency for judgment of specific traits to be influenced by the general impression of the person or thing being rated (84, p. 348).
 - b. Constant errors or systematic biases of the judges, that is, the tendency of some judges to rate uniformly high or low (84, p. 349; 67, pp. 314 ff.).
 - c. The effect of too short or too long acquaintance of the rater with the person being rated (84, p. 348).
 - d. Differences in the degree of discrimination used by different judges (67, p. 315).

46 CHECK LISTS FOR EDUCATIONAL RESEARCH

- e. The use of spurious criteria by the judges, with the result that they are not actually rating the trait which they are supposed to be rating (67, p. 317).
166. The research worker should endeavor to determine whether the proposed judgments or ratings will be sufficiently reliable to warrant their use (6, Chap. 10; 41, Chap. 9; 82; 84, p. 350). Research has shown that the average reliability of ratings on personality traits is 0.55 (84, p. 350).
167. If judgments or ratings are to be obtained, the research worker must determine what method shall be chosen. The following have been frequently used (6, pp. 122 ff.; 84, Chap. 16) :
 - The graphic rating scale.
 - The order of merit ranking.
 - The man-to-man comparison scale.
 - The method of paired comparisons.
 - The scale which provides a range of values.
168. The scale or method employed should be as simple as possible.
169. The method should be convenient.
170. The time required for recording ratings should be reasonable.
171. The scale may be devised to permit transmutation of ratings into quantitative units (84, p. 346).
172. The method of recording ratings should be as objective as possible. Placing a check is preferable to writing a word or phrase (84, p. 344).
173. Traits or terms used should be defined as objectively as possible (84, p. 352).
174. Ratings should be based on past or present accomplishment whenever possible.

175. The scale should, ordinarily, yield a normal distribution (92).
176. Directions to judges should be clear and specific (6, pp. 138 ff.).
177. The standards to be used in making judgments should be clearly defined (6, pp. 122 ff.).
178. The mechanics of recording judgments should be such as to give judges the best possible opportunity to make honest estimates (6, pp. 122 ff.; 84, pp. 338-47).
179. When human characteristics are to be rated, the judges should be directed to rate one trait at a time, disregarding every other trait (84, pp. 344, 348). It is, therefore, usually a more accurate procedure to rate a group of subjects on one trait at a time than to rate one subject on all the traits at one time (92; 84, pp. 345 f.).
180. Self-ratings tend to be too high on desirable traits and too low on undesirable traits (84, p. 349), but they are more reliable than ratings by others (84, p. 352).
181. General all-round value is often rated more reliably than specific qualities (92).
182. Analysis of the trait to be rated into sub-traits increases the number of judgments, and the reliability of the judgment, roughly to the same extent as a like increase in the number of judges, and to the extent predicted by the Spearman-Brown formula (39; 84, p. 336).
183. In constructing scales for rating human qualities or characteristics, the optimum number of class intervals or steps on the scale appears to be seven (82). If the traits or qualities are obscure and abstruse, or if the raters are untrained or

48 CHECK LISTS FOR EDUCATIONAL RESEARCH

- careless, less than seven classes are sufficient (84, p. 341).
184. In constructing scales for rating things which may be judged more objectively than human characteristics, more than seven classes should be used (84, p. 341).
185. The reliability of a rating scale is affected by the number of class intervals (82).
186. The reliability of ratings may be checked by statistical methods; for example, by correlation between ratings on the same man by the same judge at different times, by correlation between ratings on the same man by different judges, and the like (6, pp. 132 ff.; 69).
187. The research worker should note that a group of ratings may have reliability but lack validity. The agreement between the judges may be due to the fact that they are rating the same trait, but that trait may not be the one which they are supposed to be rating (67, p. 317).
188. Ranking, that is, placing individuals or things in order, is preferable to rating when groups are small, but it is no more reliable than rating (84, pp. 338 ff.).

THE JUDGES OR RATERS

189. When there are a number of available judges, the research worker should set up criteria as a basis for selecting competent judges.
190. If human characteristics are to be rated, the research worker, in order to assure fair reliability, must obtain at least three independent ratings on an objective scale by persons thoroughly acquainted with the person rated (69). For satis-

factory individual diagnosis, the average number of independent ratings needed is eight (84, pp. 353 f.).

191. In selecting judges, it should be remembered that people differ markedly in their ability to make ratings (92).
192. Judges having a common form of contact with the individual being rated (as, for example, teachers) tend to agree more closely than do judges with diverse contacts (92).
193. The average rating of a number of judges is superior to that of a single judge, provided there are not great differences in the ability of the judges (92).
194. The average rating of a small number of competent judges, for example, close associates, is likely to prove superior to the average rating of a large number of casual acquaintances (67, p. 314; 84, p. 348).
195. Long and intimate acquaintance of a judge with the person being rated tends to make ratings more favorable, that is, desirable traits are over-rated, and undesirable traits are under-rated (84, p. 349).
196. Long and intimate acquaintance of a judge with the person being rated also tends to make ratings less analytical, that is, the influence of the "halo" of general opinion is increased (84, p. 349).
197. If the rating scale is in the hands of the judges for some time prior to the date of rating, the judges have an opportunity to observe the persons to be rated, giving special attention to the traits to be rated (84, pp. 335 f.).

OTHER TECHNIQUES OF COLLECTING DATA

The above list does not exhaust the techniques of collecting data. Techniques found in special fields, for example, educational administration (75), have been omitted. A few of these methods of collecting data will be listed to indicate their specialized character. Strictly speaking, most of these are specialized, rather than additional techniques. For example, the score cards are similar to questionnaires.

198. The school survey (79; 80; 81). In the school survey a variety of techniques is employed. Procedures and techniques differ as the situation may require (3).
199. The score cards for rating buildings (76; 77; 78).
200. The score card for rating service features of a school building (26).
201. The checking list for the evaluation of supervisory and instructional programs (25).
202. The analysis of textbooks (53, p. 29).
203. The analysis of school records, board of education minutes, statutes, court decisions, etc. (53, pp. 29, 33).
204. The analysis of pupil performances, including examination papers, compositions, and the like (53, p. 29).
205. The analysis of classroom procedure, study habits, and other activities of children and teachers by the method of observation and stenographic report (53, p. 36).

CHECKING THE DATA

206. Data in the form of test scores, especially when tests have been given by assistants, should be

carefully reviewed, and part or all of the test papers should be checked.

207. If a log has been kept, this can be checked for unusual or abnormal conditions.
208. Data obtained from questionnaires, letters, interviews, etc., should be checked to discover misstatements, misinterpretations of questions, unreasonable statements, ambiguous statements, contradictory statements, and inconsistencies (70, p. 47).
209. Questionnaires should be checked as to completeness. Omissions in responses, if occurring frequently, raise a question as to the inclusion of certain items (19, pp. 196 f.).
210. Data from secondary sources should be checked to discover changes, revisions, or omissions during republication (70, p. 51).
211. The research worker should check the acceptable data against the original outline of data required for the research. He should determine whether additional data are needed (70, p. 49).

CHAPTER V

STATISTICAL METHODS USED IN RESEARCH ¹

1. The research worker should make a careful analysis of the data, looking for suggestions of meanings, inferences, and relationships. Reflective thinking must precede and accompany statistical treatment. One must literally live with the data in order to see all the implications and possibilities (70, pp. 52 ff.).
2. The research worker should recast the data in various forms, that is, apply several types of statistical treatment as outlined below, in order to derive all possible inferences (50, p. 549; 70, p. 53).
3. In connection with statistical treatment, the investigator should scrutinize and analyze the data from the qualitative point of view. For example, he should study cases that apparently fail to conform to the general trend of the results (94, p. 256).
4. In all statistical treatment, the research worker should prepare tables and graphs in the form required for publication (53, pp. 69 ff.).
5. The number of significant figures retained in all computations should correspond to the degree of accuracy and refinement of the original data (41, p. 170; 17, pp. 1-22).

¹ Although references are given to only a few books on statistical method, the research worker will find that many of the topics are treated by all standard works on statistical method.

6. The research worker should determine the appropriateness of each statistical procedure in the light of its fundamental assumptions. In all statistical treatment of data and in the interpretation of results, the investigator must constantly guard against assumptions that are only approximately true (13, p. 34).

DISTRIBUTIONS

7. The research worker may tabulate data in the form of a distribution, constructing an order distribution, a rank distribution, or a frequency distribution, as the data may require (45, pp. 354-64). Small irregularities in the distribution may be smoothed out by coarser grouping, or by the use of moving averages (40, pp. 27-31; 50, p. 260).
8. The research worker should determine the number and size of the class intervals on the basis of the nature of the data and the nature of the evidence which they are to reveal (29, pp. 1 ff.; 45, pp. 361 ff.).
9. The research worker should define the initial point of the step (29, pp. 5 f.).
10. If the statistical procedures to be used rest upon the assumption of a normal distribution, the research worker must be certain that the assumption is justified (40, p. 99; 29, pp. 81 ff.).
11. Relationships between distributions may sometimes be discovered by plotting a scatter diagram (29, pp. 153 f.).

MEASURES OF CENTRAL TENDENCY

12. When data are in the form of a distribution, the research worker may compute one or more meas-

54 CHECK LISTS FOR EDUCATIONAL RESEARCH

ures of central tendency, such as the mean, the median, or the mode² (29, Chap. 1; 40, Chap. 3; 68, Chap. 5).

13. In special cases there are other measures of central tendency which ought to be used (68, p. 244).
14. The research worker may compute percentiles, quartiles, or deciles as required (29, pp. 45 f.).

MEASURES OF VARIABILITY

15. The research worker may compute one or more measures of variability, such as range, quartile deviation, average deviation, or standard deviation (29, Chap. 1). There are other measures of variability, but they are rarely used.
16. The range, that is, the interval between the largest and the smallest measures, should be used when it is necessary to present the entire spread of the distribution or when the data are too scant to justify the calculation of another measure (29, pp. 50 f.).
17. The quartile deviation (Q) should be used when it is not desired to give undue weight to extreme scores, when only the concentration around the central tendency is sought, or when an easily computed measure is desired (29, pp. 51 f.).

² The *arithmetic mean* should be computed—

1. When it is desired to give every score equal weight in accordance with its size.
2. When the most reliable measure is desired.
3. When product-moment coefficients of correlation will be computed.
4. When standard deviations will be computed.

The *median* should be computed—

1. When the distribution is skewed to such a degree that the extreme measures would affect the mean disproportionately.
2. When a measure that can be computed easily and quickly is desired.

The *mode* should be computed—

1. When a quick approximate measure is desired.
2. When the most frequent score is desired.

18. The average deviation (A.D.) should be used when all deviations are to be weighted according to their size and when extreme deviations should not influence the measure of variability (29, pp. 51 f.).
19. The standard deviation (S.D.) or sigma should be computed when the most reliable measure is desired, when every score should have its appropriate weight, and when coefficients of correlation or measures of reliability are to be computed later (29, pp. 51 f.).
20. If coarse grouping of the data is unavoidable, the research worker should apply Sheppard's correction to the standard deviation (40, p. 167).
21. When it is desired to compare the variability of one group on two different tests, or of two groups on the same test, the coefficient of variation is sometimes used (29, pp. 40 ff.). It does not, however, seem justifiable to assume that the coefficient of variation is constant in all cases (55, p. 144).
22. It is possible to compute measures of variability or central tendency for a composite group, when the measures of variability and central tendency of each of the component groups are known (89).

CORRELATION

23. The research worker may compute coefficients of correlation when measures of relationship are desired. Since there are several methods of computing coefficients of correlation, he should understand the conditions under which each method should be used, together with the underlying assumptions (29, Chap. 4, 5; 40, Chap. 8, 9, 10, 11; 23, p. 173).

24. The most common methods of computing coefficients of correlation are:

Product-moment method, including partial³ and multiple correlation (29, pp. 163 ff.; 83).

Rank method (29, pp. 190 ff.).

η or ratio of correlation method (29, pp. 203 ff.).

Bi-serial r method (40, pp. 245-49).

Contingency method (100, p. 64; 29, pp. 195 ff.).

25. If it is difficult to tell from the appearance of a correlation table whether the regression is linear or non-linear, it is best to calculate both r and η (29, pp. 209 f.).
26. It is often desirable to reproduce correlation tables or scatter diagrams in the report of the research.
27. It is sometimes desirable to correct the coefficient of correlation for attenuation, that is, for a decreased value due to chance errors or errors of observation (29, pp. 211 f.). This correction must be properly used.
28. The coefficient of correlation may be corrected for coarseness of grouping of the data (40, p. 168).
29. A coefficient of correlation must be interpreted in the light of the variability of the group, since correlation is a function of range. Hence coefficients of correlation for populations of different ranges are not comparable (13, p. 19).
30. Coefficients of correlation may be spuriously high on account of the heterogeneity of the group (29, pp. 258 ff.).
31. If a coefficient of correlation has been calculated

³ For special formulas and illustrations of the use of partial correlation, See Franzen, Raymond, "A Comment on Partial Correlation." *Journal of Educational Psychology*, 19; 194-97, March 1928.

from data derived from a group of a certain heterogeneity, it is possible to determine the approximate coefficient in the case of a group of different heterogeneity (40, pp. 221 ff.).

32. A coefficient of correlation is affected by the selection of subjects, that is, whether they are selected from the middle or from either extreme of a distribution (34, p. 23).
33. The correlation of indices (for example, I.Q.'s) gives spurious results (29, p. 260; 13, pp. 20 f.). See special formulas (37).
34. The correlation between scores on two tests may be spurious if each test correlates with age and if persons covering a considerable span of ages are tested (13, pp. 20 f.).
35. The correlation of a test with a composite test of which it is a part yields a slightly spurious result (29, p. 260).
36. A coefficient of correlation may be interpreted in the light of its coefficient of alienation (k) (29, p. 289).
37. A coefficient of correlation must always be interpreted in the light of its probable error, and should not be reported without it (29, p. 170).
38. Coefficients of correlation are spuriously high when average scores are correlated for a number of groups, each one of which contains a large number of individuals (13, p. 23).
39. Partial and multiple coefficients of correlation must be interpreted with caution. In studies of causation it is not valid to "partial out" factors⁴ which may in part or in whole be caused by either of the two factors whose true relationship is to be measured, or by still other remote causes which

⁴ *Ibid.*

58 CHECK LISTS FOR EDUCATIONAL RESEARCH

- also affect either of the two isolated factors (12; 13, pp. 11 ff.).
40. The method of overlapping (85, p. 128) has been used to supplement the bi-serial r method.
41. When prediction is the aim, regression equations are desirable (29, pp. 173 ff.).
42. In partial regression equations, the regression coefficients cannot be interpreted as meaning that the variables in question contribute to the criterion directly in proportion to the regression coefficients (13, pp. 14 ff.).

RELIABILITY

43. Before computing measures of reliability, the research worker must be sure that the assumptions underlying the formulas are justified in each instance (29, pp. 142 ff.).
44. The research worker should compute measures of reliability, wherever needed, such as:
- a. Probable error (or standard error) of a score (40, p. 215). In comparing the probable errors of tests scaled to different units, Kelley (40, pp. 214 ff.) has suggested that the probable error of an estimated true score be divided by the standard deviation. For illustrations of this procedure, see reference 67 (pp. 74, 79, 373).
 - b. Probable error (or standard error) of a mean (29, p. 120).
 - c. Probable error (or standard error) of a median (29, p. 126).
 - d. Probable error (or standard error) of a percentile (40, pp. 86 ff.).
 - e. Probable error (or standard error) of a standard deviation (29, p. 127).

- f.* Probable error (or standard error) of Q (29, p. 128).
 - g.* Probable error (or standard error) of r (coefficient of correlation) (29, p. 170).
 - h.* Probable error (or standard error) of r that has been corrected for attenuation (40, p. 208).
 - i.* Probable error (or standard error) of estimate (29, p. 184).
 - j.* Index of reliability (40, p. 273).
 - k.* Coefficient of reliability (40, p. 200).
 - l.* Probable error (or standard error) of percentage measures of overlapping (40, p. 316).
 - m.* Correction of a reliability coefficient for curtailment of range (40, p. 223).
 - n.* Probable error (or standard error) of a difference between two averages (40, p. 182).
 - o.* Chances that an obtained difference is a true difference (29, p. 134). See also the critical ratio (49), and the experimental coefficient (46, pp. 154 ff.).
 - p.* Probable error (or standard error) of the difference between the mean of a sub-group and the mean of the total group (6, p. 172).
 - q.* Probable error (or standard error) of a coefficient of correlation estimated by means of the Spearman-Brown formula (39; 72; 73).
45. If it is necessary that a test be of a certain degree of reliability, the Spearman-Brown prophecy formula (29, p. 269) will show how much it needs to be lengthened, provided the fundamental assumptions of equal correlations and equal standard deviations are justified. However, the research worker should note that the Spearman-Brown formula does not give satisfactory prediction when the test

60 CHECK LISTS FOR EDUCATIONAL RESEARCH

material is not homogeneous, and when it is unevenly graduated (39).

46. If ratings or judgments of a certain degree of reliability are needed, the Spearman-Brown formula will show how many judges are needed (29, p. 269).
47. If the distribution is definitely asymmetrical, that fact should be taken into account in the interpretation of the probable error.

RELIABILITY OF TESTS

48. The reliability of tests should be computed for the group concerned at the time of using the test, since conditions of administration may reduce the reliability (86).
49. Reliability formulas measure the unreliability due to errors which arise from sampling, but not the unreliability due to other causes, such as an inadequate number of tests, chance errors, and constant errors (8; 29, p. 143).
50. In reporting the reliability of tests, complete data should be given (65; 67, p. 365), such as
 - a. The reliability coefficient (together with a statement of the method of computation; for example, correlation of chance halves, correlation of scores obtained by repetition of the same form, or correlation of similar forms).
 - b. The number of cases used.
 - c. The standard deviations of the test form or forms.
 - d. The probable error of a raw score.
 - e. The probable error of an estimated true score.
 - f. The probable error of the estimated true score divided by the standard deviation.

- g.* The nature of the group (range of grades or ages, or other data showing the representative character of the sampling).

MISCELLANEOUS

51. The statistical treatment of the data should be of a degree of refinement that is in harmony with the character and accuracy of the data (18).
52. Half-truths must not be accepted as whole-truths simply because they seem to be supported by statistical evidence (16, p. 17).
53. The statistical treatment should be adequate.
54. Special or new uses of statistical formulas should be fully explained and substantiated.
55. The derivation and substantiation of new statistical methods should be given in full.
56. All mathematical and statistical computations must be checked to insure accuracy.
57. When percentages are given, the absolute numbers should also be reported.
58. Due allowance should be made for all kinds of errors, especially when measurements are employed; namely (36),

Scale errors, that is, defects of the test or scale.

Sampling errors.

Scoring errors.

Response errors; for example, the subjects may respond differently on successive trials under differing circumstances.

59. When scores on tests are to be combined, the method proposed should be scrutinized in the light of underlying assumptions (6; 29, pp. 277 ff.).

62 CHECK LISTS FOR EDUCATIONAL RESEARCH

60. Scores may be converted to standard measures (40, p. 115).
61. Test scores may be T scaled (46, pp. 98 ff.).
62. The percentage of a group choosing one of two alternatives may be converted into the sigma value of such a percentage (85, p. 115).
63. Index numbers may be used to advantage in cost studies (40, p. 331; 50, Chap. 6, 9), and in studies comparing cities, states, etc. (5; 10; 11; 27).
64. Special formulas, such as Mort's for the measurement of educational need, may be obtained by reference to the literature of the particular field.
65. Measures in the form of ranks, that is, relative positions, may be transmuted into units of amount (29, p. 111).
66. To make allowance for the fact that some judges rate consistently higher than others, statistical treatment must be employed (6, p. 140).
67. If incomplete "order of merit" ratings are to be combined, the research worker should make a choice from such methods as the following (30):
 - Averaging the ranks available.
 - Percentile method.
 - Sigma position method.
68. When data are in the form of ranks (for example, when a teacher has ranked the members of a class), it is possible to transmute these ranks into approximate age or grade scores by the use of a special technique (45, pp. 60 ff.).
69. Ranks may be transmuted into sigma positions (21, p. 13).

CHAPTER VI

THE PRESENTATION OF THE RESEARCH

It is assumed that the research worker has completed his investigation and the statistical treatment of his results. He is now ready to report his findings. The following Check Lists are based on the assumption that the research is of such proportions as to warrant presentation in book or pamphlet form. If this is not the case, the presentation may be simplified accordingly.

The report of a research should include at least four divisions, but the chapters or sections may be arranged in accordance with the requirements of the particular research. These major divisions are:

1. Introduction.
2. Statement and definition of the problem.
3. Presentation of the data.
4. Conclusion.

INTRODUCTION

5. An introduction, foreword, or preface should be prepared (53, p. 62). In reporting researches, these terms may be considered synonymous (70, p. 72).
6. The research worker should include in the introduction explanatory statements that will acquaint the reader with the problem under consideration (53, p. 62).
7. The introduction may include a statement of the

64 CHECK LISTS FOR EDUCATIONAL RESEARCH

importance of the problem, the circumstances leading to the research, and the purpose in the mind of the writer (53, pp. 58 ff.).

8. The research worker frequently wishes to acknowledge assistance rendered by other persons or the use of quotations from books or manuscripts. Such acknowledgments may be included in the introduction or preface.
9. If the research is published under the editorship of some one other than the author, there will usually be an Editor's Introduction in addition.

STATEMENT AND DEFINITION OF THE PROBLEM

10. The problem should be precisely defined and fully described either in the introduction or in the first chapter (53, Chap. 2). The research worker should inform the reader just what questions are to be answered.
11. The research worker should show the relation of the problem to other problems and questions.
12. The problem should be defined in such a way that the reader will understand the probable applications of the findings to his own problems and interests (53, p. 58).
13. The writer may deem it desirable to show the worthwhileness of the study (53, p. 25).
14. The objectives of the research should be clearly set forth, together with the subordinate questions to be answered, and the limits of the investigation (53, pp. 14 ff.).
15. The character of the data which it is proposed to assemble should be described.
16. A statement indicating the limitations of the data or techniques may be necessary (53, p. 24).

17. In the light of its possible bearings on the conclusions, the writer may wish to show that the data are comprehensive.
18. If the data are representative rather than comprehensive, the writer should present all essential information.
19. When subjects have been chosen for testing or experimental purposes, the method of selection should be set forth (Chap. III).
20. If the problem has been previously studied by other research workers, the writer should present a critical review of these studies, possibly with citations. The relation of this study to the others should be clear (53, pp. 22 f.).
21. Terms should be defined, particularly those that are technical or are used in a technical sense, as well as others that may be unfamiliar to the reader (53, p. 25).
22. Assumptions upon which the research is based and their implications should be explicitly stated (53, p. 25).
23. The research worker will sometimes find it necessary to state the point of view or philosophy of education upon which the investigation is based.
24. A sufficient number of pertinent details should be presented so that the research may be repeated or checked by others.

PRESENTATION OF DATA

25. In writing the report, the research worker should aim to make the reader understand the data that are presented.
26. He should review, mentally, the successive stages in the collection and treatment of data to deter-

66 CHECK LISTS FOR EDUCATIONAL RESEARCH

mine how much space should be given to a description of the method of collecting the data, and to decide what data are essential and desirable in the presentation (53, Chap. 6).

27. In describing the method of collecting data, the chronological order will frequently prove to be the clearest and most desirable arrangement.
28. Much or all of the original data may be presented, especially if the field is comparatively undeveloped, or if novel methods have been employed (70, p. 55).
29. If the data in the original form are cumbersome, excerpts may be presented or the data may be placed in an appendix (70, pp. 59 f.).
30. If the data in the original form will require more time and effort to comprehend than seems justified, a description may be given instead (70, p. 56).
31. If the treatment of the data has been simple and straightforward, the final form of the data may be presented, together with an explanation of the treatment (70, pp. 57 ff.).
32. If such instruments as tests, questionnaires, rating scales, and the like are too cumbersome for presentation in the text, they may be placed in the appendix.
33. The form of the data should be such as to set forth the evidence in unmistakable terms (70, p. 58).
34. Data or explanations that would interrupt the sequence of the reader's thought should be placed in footnotes or appendices (70, p. 60).
35. The research worker should evaluate the data and his interpretation so as to omit irrelevant and unnecessary data and ideas (53, pp. 63 f.).

36. Under some circumstances it is desirable to report methods of attack or treatment that proved fruitless, especially if this will be a valuable contribution to future workers.
37. Unusual or complex techniques should be fully explained.
38. The presentation as a whole should be logical, well organized, and coherent (53, pp. 62 ff.).
39. The research worker should strive for precision of expression by wise choice of words and by skillful construction of sentences (53, pp. 64 ff.).
40. The presentation should be simple and direct, and the writer should avoid gaps in the trend of thought (53, pp. 62 f.).
41. The research worker should strive for clearness by using a suitable vocabulary and simple, definite statements (53, pp. 66 ff.).
42. Summaries at the end (or beginning) of each unit of the research are desirable. These summaries should include the high lights or significant points, so arranged as to assist the reader in organizing his knowledge.
43. The research worker should be careful to follow the accepted principles of rhetoric, grammar, spelling, and punctuation (53, pp. 68 ff.).
44. In short, the exposition should be such that the average student in the field may be able to follow the presentation of data with facility and interest.

TABLES

45. Tables should be so placed that interpretation both of text and of tables is facilitated.
46. Tables should be arranged to read from left to right.

68 CHECK LISTS FOR EDUCATIONAL RESEARCH

47. Every table should be given a number, and Roman numerals are most frequently employed for this purpose (53, pp. 70 f.).
48. Tables should be numbered consecutively throughout the research.
49. Tables should usually be referred to by number; references such as "the above table" and "the following table" should be avoided because it is not always possible on the printed page to place a table immediately above or below the reference to it in the text.
50. Every table should have a caption or title for the purpose of indicating its contents. It should be placed above the table.
51. The number of the table, preceded by the word "Table," may be printed above the caption or on the same line as the caption, immediately preceding it.
52. Periods are not usually used in titles or in column headings, except at the end of a sentence which does not complete the title.
53. In a column of financial data, the dollar sign and the per cent sign should appear only with the first entry and before the total.
54. Tables should be as simple as the nature of the data permits, including only so much data as are pertinent and essential (53, pp. 46 f., 69 f.).
55. Tables should be as short as possible, other things being equal, and should be confined to a single page when practicable. Large tables may often be broken into two or more separate tables.
56. Rules, both vertical and horizontal, should be used in tabular matter to separate headings and text when spacing alone is not sufficient (53, pp. 70 f.).

57. If it is necessary to print complex tables of original data of minor importance, they should be placed in an appendix.
58. Tables should have captions that can be readily understood without referring to the text for an explanation (53, p. 69).
59. When different groups are being studied, it is well to designate each one by name or number to facilitate references to the tables.
60. The writer should organize his material so as to fit the length and breadth of the table to the size of page of the publication in which it will appear (53, pp. 46 f.).
61. The text accompanying a table or a graph should be so complete that it will be unnecessary for the reader to refer to the table or graph in order to follow the trend of thought (53, p. 72).

GRAPHS AND CHARTS

62. Graphic presentation may be used for data already presented in tables in order to clarify the exposition (95; 40, p. 9).
63. Graphs should be numbered consecutively, the common practice being to use Arabic numerals (53, p. 72).
64. Every graph should have a caption, that is, a title or heading placed, preferably, at the bottom, but such phrases as "graph showing" and all unnecessary words should be avoided. Captions of similar graphs should have similar wording.
65. Graphs should usually be referred to by number.
66. Graphs should be so simple that they can be understood at a glance (45, Chap. 13).

70 CHECK LISTS FOR EDUCATIONAL RESEARCH

67. "The general arrangement of a diagram should proceed from left to right¹ (62; 45, p. 334).
68. "Where possible, represent quantities by linear magnitude, as areas or volumes are more likely to be misinterpreted (62; 45, p. 334).
69. "For a curve, the vertical scale, whenever practicable, should be so selected that the zero line will appear in the diagram (62; 45, p. 334).
70. "If the zero line of the vertical scale will not normally appear in the curve diagram, the zero line should be shown by the use of a horizontal break in the diagram (62; 45, p. 335).
71. "The zero lines of the scales for a curve should be sharply distinguished from the other coördinate lines (62; 45, p. 335).
72. "For curves having a scale representing percentages, it is usually desirable to emphasize in some distinctive way the 100 per cent line used as a basis of comparison (62; 45, p. 336).
73. "When the scale of the diagram refers to dates, and the period represented is not a complete unit, it is better not to emphasize the first and last ordinates, since such a diagram does not represent the beginning and end of time (62; 45, p. 336).
74. "When curves are drawn on logarithmic coördinates, the limiting lines of the diagram should each be of some power of 10 on the logarithmic scale (62; 45, p. 337).

¹ This and the following items in this section are quoted from the *Report of the Joint Committee of the American Society of Mechanical Engineers on Standards for Graphic Presentation* (62). This report is quoted, with illustrations, in reference 45, Chap. 13. For further principles of graphing see reference 45, pp. 341-53. A checking list for graphic presentations, by W. C. Brinton, may be found in his *Graphic Methods for Presenting Facts*. Engineering Magazine Company, 1914. It is reproduced in reference 95, pp. 314 f.

75. "It is advisable not to show any more coördinate lines than necessary to guide the eye in reading the diagram (62; 45, p. 337).
76. "The curve lines of a diagram should be sharply distinguished from the ruling (62; 45, p. 337).
77. "In curves representing a series of observations, it is advisable whenever possible to indicate clearly on the diagram all the points representing the separate observations (62; 45, p. 338).
78. "The horizontal scale for curves should usually read from left to right and the vertical scale from bottom to top (62; 45, p. 338).
79. "Figures for the scale of a diagram should be placed at the left and at the bottom or along the respective axes (62; 45, p. 339).
80. "It is often desirable to include in the diagram the numerical data or formulae represented (62; 45, p. 340).
81. "If numerical data are not included in the diagram it is desirable to give the data in tabular form accompanying the diagram (62; 45, p. 340).
82. "All lettering and all figures in a diagram should be placed so as to be easily read from the base as the bottom, or from the right-hand edge of the diagram as the bottom (62; 45, p. 341).
83. "The title of a diagram should be made as clear and complete as possible. Sub-titles or descriptions should be added if necessary to insure clearness" (62; 45, p. 341).

CITATIONS

84. Citations may be used to orient the research with respect to allied studies (70, p. 62).
85. Citations may be used in support of data, methods,

72 CHECK LISTS FOR EDUCATIONAL RESEARCH

or techniques used by the research worker (70, p. 63).

86. Citations may be used to show the concurrence or non-concurrence of others in the inferences, opinions, or conclusions of the writer (70, pp. 63 f.).
87. Citations from laws, official documents, and the like, are commonly made for the purpose of retaining the exact words of the original (70, p. 66).
88. Citations should be exact as to spelling, phraseology, and the like.
89. Consent of the author should be secured if extensive citations are made.
90. Citations should be accompanied, in either text, footnote, or bibliography, by the author's name, the exact title of the work, publisher, place and date of publication, total number of pages, and page or pages cited. This information should be given in an approved form, which will conform to the practice adopted by the publisher.

REFERENCES

91. References are employed to refer to the work of others, when citations are unnecessary or cumbersome (53, pp. 72 ff.).
92. References may be used for all purposes for which citations are used.
93. References may be used to give the sources of data (70, p. 67).
94. References may be used to refer to other parts of the research, and to make clear relations between the discussion and the data (70, p. 67).
95. References should always include complete bibliographical data, unless these data are given in the bibliography (53, pp. 72 ff.).

CONCLUSION

96. It is usually desirable to devote a chapter or a section to summarizing the research and the conclusions (53, Chap. 5).
97. The summary of methods, techniques, and treatment of data should be concise yet complete enough to enable the casual reader to determine whether he wishes to read the entire study (53, p. 50).
98. The conclusions should follow logically from the evidence submitted, and should be specifically connected with the questions in the original statement of the problem (53, p. 50).
99. The conclusions should be warranted by the data presented, any limitations due to incomplete data being duly recognized (53, p. 52).
100. The conclusions should be free from bias.
101. The research worker must show that he has properly analyzed and evaluated the data by indicating on what findings or assumptions each conclusion is based.
102. Conclusions should not be based on assumptions unless the assumptions have been justified (53, p. 25).
103. The research worker should make clear to what extent conclusions inevitably follow, and to what extent they are matters of interpretation (53, p. 52).
104. Conclusions should be clearly and succinctly stated.
105. Conclusions should be stated so simply that the reader will understand the meaning and implications of the results.

74 CHECK LISTS FOR EDUCATIONAL RESEARCH

106. The application of the conclusions to educational practice should be expressed in non-technical language which will be intelligible to those who may not understand the techniques employed (53, p. 53).
107. Conclusions should be so stated that a reader can secure the essential results of the study by reading only the conclusions.
108. In stating the conclusions, the research worker should indicate his contribution, that is, wherein he corroborates or disproves existing studies, and wherein he has revealed additional facts.
109. The research worker should indicate whether the research has been sufficiently extensive to merit general application of the findings (53, p. 52).
110. If the findings are not capable of general application, he should specify the limits of age, grade, sex, mental ability, or the like to which the conclusions apply (53, p. 53).
111. The final or tentative character of the findings should be clearly shown (53, p. 52).
112. The research worker should set forth the negative as well as the positive results.
113. Exceptions to principles should be discussed for what they are worth.
114. Artificial, atypical, or unusual conditions in the course of the research should be given due consideration in deriving the conclusions.
115. Due attention should be given to valuable conclusions in the way of by-products.
116. The writer should check against the problems, especially the sub-problems raised at the beginning of the research, to make sure that they have been included in the conclusions.

117. Suggestions for further research may be made (53, p. 54).
118. The research study as a whole should give to the reader an impression of unity (53, Chap. 6).
119. The writer may present recommendations in the light of the conclusions and findings (53, p. 54).

THE BIBLIOGRAPHY

120. Except in rare instances a bibliography should be appended to the research. The bibliography should include complete information; namely, author's name, the exact title, publisher, date of publication, and total number of pages. If the publisher is comparatively unknown, the place of publication is also needed. For periodical references, the following items should be given: author, title, name of periodical, number of volume, page numbers on which article appears, month and year of publication. Other types of references have special forms (53, pp. 73 ff.).
121. The bibliography should contain the references which the writer has found helpful.
122. The bibliography should contain all works from which citations have been taken, or to which reference has been made.
123. The references in the bibliography should usually be listed alphabetically by authors' names (53, p. 73).
124. The references in the bibliography may be classified on the basis of content, the references within each group being listed alphabetically.
125. The annotated bibliography is obviously more useful to the careful reader (60, p. 81).
126. Important references may be starred.

THE APPENDIX

127. Materials too voluminous for footnotes may be placed in an appendix (70, p. 75).
128. It is often desirable to place unimportant original data in the appendix.
129. Forms, questionnaires, and other material not relevant to the body of the text should be placed in the appendix.
130. The material in the appendix must be homogeneous; if not, it should be segregated into several appendices. When there are several appendices, they are usually headed Appendix A, Appendix B, etc.

THE FORM OF PRESENTATION

131. The title should be well chosen, challenging to the reader, and accurately descriptive of the problem. The form of the title page should conform to the usual practice of the publisher (53, p. 82).
132. There should be a table of contents which should contain the titles of all chapters, and which may include the important subdivisions, thus giving a definite idea of the scope of the subjects presented (53, pp. 76, 83).
133. There should be a list of tables, giving the captions of the tables, preceded by the Roman numerals. If the research includes graphs, illustrations, or figures, separate lists should be given for each. The titles should be preceded by the Arabic numerals (53, p. 83).
134. There should be chapter titles or headings which are indicative of the content of each chapter (53, p. 76).

135. Section or paragraph headings will prove a valuable aid to the reader (53, p. 79).
136. Citations should be clearly indicated as such, in one of the following ways: indentation, single spacing, smaller type, or quotation marks (70, p. 65).
137. Sources of citations may be stated before or after the quotation, or in a footnote (70, p. 65).
138. Elisions, that is, omissions, should be indicated by ellipses or periods, thus . . ., or by the phrase "and so forth." When a whole paragraph or a series of paragraphs is omitted from a citation, a full line of periods is usually inserted to indicate the omission (70, p. 65).
139. Corrections, remarks, or words inserted by the one who makes a citation are usually placed within brackets [].
140. In a typewritten manuscript, footnotes should be inserted on a line immediately following that in which the reference is made. The footnote should be separated from the text by horizontal lines. If the manuscript is not intended for publication, footnotes are usually placed at the bottom of the page (53, p. 80).
141. Footnotes should always give enough specific information to enable the reader to locate the reference easily, exact page numbers being desirable (53, p. 76). If there is a final bibliography, however, data on a reference other than author, title, and page numbers are not needed in footnotes.
142. Uniformity should be observed in regard to the punctuation, capitalization, and order of entry of the data in footnotes.

78 CHECK LISTS FOR EDUCATIONAL RESEARCH

143. Figures are preferable as index signs to footnotes. They may run consecutively for a page, a chapter, or the entire study (70, p. 70).
144. In tables, especially those in which figures are used, lower case letters or reference symbols such as asterisks, daggers, etc., are preferable as index signs (70, p. 70).
145. Each footnote should be paragraphed.
146. Abbreviations may be used in making cross references in footnotes; for example,
 - Ibid.* (meaning *the same*) may be used when reference is made to the source cited immediately above.
 - Op. cit.* (meaning *the work cited*) may be used when other references have intervened. It is usually preceded by the author's name.
 - Loc. cit.* (meaning *the place cited*) may be used, following an author's name, to refer to the same passage of the source cited.
 - Cf. supra* is used to refer to something which precedes.
 - Cf. infra* is used to refer to something which follows.
 - Passim* is used in referring to material scattered through the source.
147. An index is frequently desirable, especially if the work is lengthy, or if references are likely to be made to particular items or sections.
148. Multiple references should be included in an index.
149. A glossary is usually needed if a number of technical, unusual, or difficult terms are employed.

ENGLISH STRUCTURE AND FORM

150. Appropriate words and phrases should be used (53, p. 66).
151. Words and phrases in common use, if used in a technical sense, should be carefully defined (53, p. 67).
152. Words should not be overworked, that is, repeatedly used to the point of monotony (53, p. 67).
153. Unity, coherence, and emphasis in paragraph construction should be observed (53, p. 68).
154. Topic sentences, placed at the beginning of a paragraph and consistently held in mind, will contribute to better unity.
155. The sequence of ideas may be checked by underlining topic sentences and reading them throughout the report, from beginning to end.
156. Grammar, spelling, and punctuation should be carefully checked for errors.
157. Statements should be worded so as to avoid ambiguity or indefiniteness (53, p. 67).

CHAPTER VII

THE APPLICATION OF CHECK LISTS IN THE ANALYSIS OF PUBLISHED RESEARCHES

THE above outlines or analyses of educational research are intended primarily as guides during the progress of a research. They have been termed Check Lists because the investigator may check his plans and procedures against the items in the various chapters. As a particular procedure or technique is completed, he may check off that particular item, and proceed to the next.

The Check Lists may also be used to analyze a completed research. An individual or a committee may desire to check a thesis, a dissertation, or other published research in order to determine whether the research is scientific. The educational world to-day finds a multitude of investigations and researches, the conclusions of which attack or defend the present curriculum, methods of teaching, administration of schools—in fact every phase of our educational program. Many are conflicting in their conclusions. Some are clearly unscientific in method of research, techniques of collecting data, or statistical methods. The careful student of education will wish to analyze many such investigations, and these Check Lists will prove to be a convenient instrument.

Another possible use of the Check Lists is in connection with the training of research workers. If a formal

course, such as a course in Methods of Research or Research in Secondary Education, is offered, the question arises as to the relative importance of the various methods and techniques suggested by the Check Lists.

In the light of the possible uses of the Check Lists, suggested above, the following questions may be raised:

1. Have these methods and techniques actually been used by research workers?
2. What is the relative frequency with which these methods and techniques have been used?
3. Are the Check Lists reasonably complete?

In order to answer these questions, a representative selection of researches was read, and checked against the outlines in the foregoing chapters.

THE SELECTION OF THE RESEARCHES

Doctoral dissertations seemed to offer the most promising field for analysis since they afford a representative group of scientific researches and indicate the probable relative frequency of use, by research workers in the future, of the methods and techniques of research.

In order to obtain a representative group of dissertations, it was planned to analyze one hundred of those which were published during the years 1924 and 1925. The titles were obtained from *A List of American Doctoral Dissertations* (91) prepared by the Library of Congress, and the *Register of Teachers College Doctors of Philosophy 1899-1925* (61). These sources may be assumed to furnish a reasonably complete and certainly a representative list. One hundred nine dissertations were located through these two sources. Three of these dissertations employed philo-

sophical techniques. For example, *Education as Growth* was one of the titles. Because of the fact that such techniques could not be included in the Check Lists, these three dissertations were eliminated. Six more dissertations were eliminated by a process of random selection. All of the nine dissertations thus eliminated were prepared under the auspices of Teachers College, Columbia University.

TABLE I

DISTRIBUTION, BY UNIVERSITIES, OF DISSERTATIONS SELECTED FOR ANALYSIS. YEARS 1924 AND 1925

UNIVERSITY	No.
Brown University	1
Columbia University	5
Johns Hopkins University	1
Leland Stanford University	2
Northwestern University	1
Ohio State University	2
Teachers College, Columbia University	73
University of Chicago	2
University of Iowa	5
University of Kansas	1
University of Michigan	2
University of Pennsylvania	5
Total	100

Table I lists the names of the universities under whose auspices the dissertations were prepared, together with the number from each institution. Attention is called to the fact that some institutions do not require that doctoral dissertations be published. In other cases, the dissertations are not published until

some time after they are accepted. These facts explain the preponderance of dissertations from Teachers College, Columbia University, and the comparatively small number from certain other well-known institutions. In no sense, therefore, may Table I be interpreted to include all the doctoral dissertations written and accepted during the years 1924 and 1925.

ANALYSIS OF THE DISSERTATIONS

The results of the analysis of these dissertations are found in Tables II, III, IV, V and VI. These tables were compiled by reading the dissertations and checking the techniques and procedures used. If, however, a certain technique, such as the product-moment correlation method, was used ten times in one dissertation it was checked only once.

TABLE II

ANALYSIS OF THE STATEMENT OF THE PROBLEM IN ONE HUNDRED DISSERTATIONS

	ELEVEN UNIVER- SITIES	TEACHERS COLLEGE COLUMBIA UNIVERSITY	TOTAL
	<i>N</i> = 27	<i>N</i> = 73	<i>N</i> = 100
Worthwhileness of the Problem	24	65	89
Survey of the Literature	19	41	60
Formulation of Objectives	23	62	85
Statement of Data Required	19	40	59
Comprehensive Character of the Data . . .	7	9	16
Representative Character of the Data . . .	13	35	48
Definition of Terms	2	12	14

THE STATEMENT OF THE PROBLEM IN ONE HUNDRED
DISSERTATIONS

Table II shows clearly that in these dissertations there was no uniform plan for the statement of the problem. Uniformity is not essential, perhaps not desirable. The application of the Check Lists, however, showed certain definite trends, as follows:

1. Eighty-nine per cent of the writers gave some space to an explanation of the worthwhileness of their problems. The remainder probably felt that the importance of the problem was self-evident.

2. Sixty per cent of the writers presented a survey of the literature. Some writers, however, were working in practically virgin soil. In a few cases a discussion of the findings of other investigators was given elsewhere in the research, when comparisons with the writer's findings were possible.

3. Eighty-five per cent of the group clearly set forth the objectives of the research and fifty-nine per cent described the data to be obtained in order to reach these objectives.

4. In some researches, the data were comprehensive; in others, representative. A majority of the writers pointed out the comprehensive or representative character of the data in connection with the statement of the problem. Others preferred to do so as the data were presented. In a few cases the data were neither comprehensive nor representative.

5. Few writers deemed it advisable or necessary to define terms in connection with the statement of the problem. Such definitions, especially of unusual statistical terms, would have made certain dissertations more readable to the uninitiated.

TABLE III

ANALYSIS OF THE METHODS OF RESEARCH USED IN ONE HUNDRED
DISSERTATIONS

	ELEVEN UNIVER- SITIES	TEACHERS COLLEGE COLUMBIA UNIVERSITY	TOTAL
	<i>N</i> = 27	<i>N</i> = 73	<i>N</i> = 100
Historical Method	6	15	21
Experimental Method			
One-group	3	9	12
Equivalent-groups	8	13	21
Rotation	0	0	0
Case Method	2	7	9
Case-Group Method	0	1	1
Investigational Methods	20	56	76
Totals	39	101	140

METHODS OF RESEARCH IN ONE HUNDRED
DISSERTATIONS

In the Check Lists we have somewhat arbitrarily made a distinction between methods of research and techniques used in research. Application of the Check Lists to one hundred dissertations revealed the fact that many dissertations employed only one method of research, while others employed two or more methods. It will be noted in Table III that one hundred forty methods of research were employed in these one hundred dissertations.

It is clear that the research worker must understand the historical method. To be sure, very few of the dis-

86 CHECK LISTS FOR EDUCATIONAL RESEARCH

sertations examined used only this method. Historical introductions were found, however, in about one-fifth of the dissertations.

The student who is inclined to think of experimentation as the tool of the exact sciences only, such as physics and chemistry, will be surprised to find that thirty-three dissertations, or one-third of the total number, found it necessary to experiment. In education, the experiment is a rather hazardous type of research, so many are the pitfalls and variable factors. Graduate students obviously need more training and practice here. Significant also is the fact that most of these experiments were conducted in classrooms or with the assistance of teachers in service.

The case method and the case-group method are employed by the educational sociologist, and to a certain extent by the psychologist. These techniques are comparatively new. As yet they can hardly be said to be as exact as the other methods described.

Investigational methods were employed by seventy-six per cent of these writers. As has already been shown, this method aims to discover factors that are already being applied or are in operation, and to study their effects. In training research workers, investigational methods should have an important place, not only because they are likely to be used by many research workers, but also because there are many different kinds of investigations.

TECHNIQUES OF COLLECTING DATA IN ONE HUNDRED DISSERTATIONS

Table IV shows that the research worker in the field of education must have at his disposal many techniques for collecting data. Thirty-eight of these research

TABLE IV

ANALYSIS OF TECHNIQUES OF COLLECTING DATA IN ONE HUNDRED DISSERTATIONS

	ELEVEN UNIVER- SITIES	TEACHERS COLLEGE COLUMBIA UNIVERSITY	TOTAL
	N = 27	N = 73	N = 100
Measurement			
Evaluation and selection of existing tests	11	27	38
Construction of new tests	10	13	23
Administration of tests	12	25	37
Scoring of tests	12	20	32
Interviews	5	11	16
Questionnaires	11	31	42
Necessity for, and value of, shown	6	16	22
Scientific selection and arrangement of items	3	3	6
The original form presented	5	15	20
Judgments or Ratings	9	14	23
Qualifications of the judges	3	8	11
Personal Observation, including Steno- graphic Reports	3	8	11

workers had to select tests from the large number of measuring instruments available. Since the conclusions of so many researches are based on the results of tests, it is clear that most research workers must be thoroughly competent in this field. The measurement movement is still young, and the novice may make serious mistakes in his choice of tests and interpretation of results.

It was somewhat of a surprise to find that practically one-fourth of this group of workers had to con-

struct their own tests. Test building has become a highly technical process. Obviously, future research workers must be trained to construct tests, at least until the supply of tests is more nearly complete than it is at present.

Most of those who employed tests described the precautions which they had used in administering and scoring them. These precautions are of vital importance, particularly if a number of examiners and scorers are employed. Complete information enables a critical reader to evaluate the results of the tests and the interpretations of the research worker.

Interviews proved to be needed in but sixteen per cent of the cases, but the questionnaire, which is a form of interview after all, was used by forty-two workers. This is a surprisingly large number in view of the criticisms and attacks on the validity of the questionnaire, and some writers did not demonstrate the necessity for its use. Only six explained their methods of selecting and arranging the items, although others may have been just as careful and scientific in preparing their questionnaires. The lack of information in regard to methods of selection and organization of items is regrettable. It is also to be regretted that only about one-half of those who used a questionnaire published the actual questionnaire in its original form. It is certainly to be recommended that all forms, blanks, questionnaires, and the like, be published, possibly in an appendix to the study.

As noted in the check list on judgments or ratings, statistical studies have shown the lack of validity and reliability in many cases where ratings have been used. We note that twenty-three of these one hundred research workers used ratings. Only eleven of the

twenty-three were careful to set forth the qualifications of the judges.

Eleven of the research workers used personal observation. This technique has its pitfalls, but seems at times to be essential.

Summarizing the techniques used in collecting data, as presented in Table IV, it would seem that thorough training in measurement is essential for prospective research workers. Training and practice in the preparation of questionnaires, and cautions as to their use should have an important place in such a course. Research workers also need to be aware of the pitfalls in the use of ratings. Rules for the construction of rating scales have been prepared by competent research workers. Finally, it would seem to be desirable to train research workers for interviewing. Although this technique may not be used in collecting data, the novice at research will need to have many interviews with the experienced research workers who are directing or assisting him in his investigations.

STATISTICAL COMPUTATIONS IN ONE HUNDRED DISSERTATIONS

Those who assert that statistical method is a more useful instrument of educational research than the foreign languages will find abundant support in Table V. It may be safely said that the great majority of these dissertations employed statistical methods.

As might have been anticipated, measures of central tendency, next to distributions, were the most frequently used measures. Although reported in only thirty-six cases, standard deviations were actually computed in obtaining a product-moment r , hence the standard deviation is really an important measure, as

90 CHECK LISTS FOR EDUCATIONAL RESEARCH

TABLE V

ANALYSIS OF THE STATISTICAL METHODS EMPLOYED IN ONE HUNDRED
DISSERTATIONS

	ELEVEN UNIVER- SITIES	TEACHERS COLLEGE COLUMBIA UNIVERSITY	TOTAL
	<i>N</i> = 27	<i>N</i> = 73	<i>N</i> = 100
Distribution	21	60	81
Measures of Central Tendency			
Mean or average	15	45	60
Median	11	41	52
Mode	4	3	7
Percentile Measures			
Quartiles (Q_1 , Q_3)	4	17	21
Percentile, decile, quintile	4	7	11
Measures of Variability			
Range	8	33	41
Standard deviation	11	24	35
Quartile deviation	3	5	8
Median deviation	0	1	1
Average deviation	0	2	2
Coefficient of variation	2	1	3
Measures of Relationship			
Product-moment r method	13	28	41
Partial correlation method	5	11	16
Multiple correlation method	3	4	7
Rank method	8	6	14
η or ratio of correlation method	0	4	4
Bi-serial r method	2	1	3
Coefficient of mean square contingency method	1	5	6
Coefficient of association method	0	1	1
Correction for attenuation	2	6	8
Correction for restricted range	0	1	1
Coefficient of alienation	0	2	2
Predictive index	0	5	5

TABLE V (Continued)

	ELEVEN UNIVER- SITIES	TEACHERS COLLEGE COLUMBIA UNIVERSITY	TOTAL
	<i>N</i> = 27	<i>N</i> = 73	<i>N</i> = 100
Measures of Reliability			
Probable error of a score	4	4	8
Probable error or standard error of a mean	8	12	20
Probable error or standard error of a median	0	1	1
Probable error of <i>r</i>	11	21	32
Probable error of <i>r</i> that has been cor- rected for attenuation	1	0	1
Probable error of estimate	1	1	2
Index of reliability	0	2	2
Probable error of a sum	0	1	1
Probable error or standard error of dif- ference	9	15	24
Chances that obtained difference is a true difference	8	11	19
Experimental coefficient	0	3	3
Spearman-Brown prophecy formula ...	4	8	12
Regression Equation	3	5	8
Miscellaneous			
T scores	0	1	1
Index numbers	1	3	4

judged by the number of computations. It is also safe to say that the typical research worker will not need to know how to compute median deviation, average deviation, or quartile deviation.

The facts concerning measures of relationship may surprise some who have been content to master only the rank and product-moment methods of computing

correlation. Sixteen workers computed partial correlations, and there is a fair sprinkling of the other, less commonly used, methods.

Measures of reliability seem to have been neglected. Sixty writers computed a mean, but only twenty reported the probable error of the mean. Fifty-two computed a median, but only one reported the probable error of the median. A large number who computed correlations failed to report the probable errors.

Keeping in mind the number of experiments previously reported, it is of interest to note that the probable error of the difference was reported in only twenty-four cases. The critical ratio, although not separately listed, was found in several of the nineteen dissertations which reported the chances that the obtained difference was a true difference. The experimental coefficient was found in three dissertations.

We may summarize this section, in conclusion, by saying that practically all prospective workers in the field of educational research will do well to equip themselves with a knowledge of elementary statistical methods. However, the data in Table V must not be used as a measure of the relative importance of the various statistical techniques. We note that only one worker used T scores and only twelve employed the Spearman-Brown prophecy formula. But all research workers who use measurement should be familiar with these two procedures. In the one hundred dissertations it is probable that writers used the techniques that they knew. For future use others may be better. Hence, in planning a course in statistical methods, one must consider not only those procedures which are frequently used, but also those which are frequently overlooked.

TABLE VI

ANALYSIS OF METHODS OF PRESENTATION USED IN ONE HUNDRED
DISSERTATIONS

	ELEVEN UNIVER- SITIES	TEACHERS COLLEGE COLUMBIA UNIVER- SITY	TOTAL
	N = 27	N = 73	N = 100
Tables	26	70	96
Graphs or Charts	14	43	57
Illustrations	7	9	16
References, Footnotes	25	72	97
Chapter Summaries	15	38	53
Final Summary	23	61	84
Bibliography	20	54	74
Annotated Bibliography	0	4	4
Table of Contents	26	72	98
Table of Tables, Charts	11	45	56
Paragraph or Section Headings	26	66	92
Index	7	10	17

METHODS OF PRESENTATION IN ONE HUNDRED
DISSERTATIONS

These one hundred writers conform very closely to the accepted methods of presentation. The large number using graphs would tend to show that some attention should be given, in a Methods of Research course, to the standards for graphic presentation.

The final summary is used by eighty-four per cent of the group. Chapter summaries are found in fifty-three per cent of the dissertations.

Of the seventy-eight bibliographies presented, only four were annotated.

Practically all the writers presented tables, but not

94 CHECK LISTS FOR EDUCATIONAL RESEARCH

all of them included a list of these tables along with the table of contents. Such a table of tables is not required in all cases, especially if the sub-topics of the text cover the table topics.

Paragraph and section headings were commonly used, and added to the lucidity of the presentation.

Throughout this discussion no effort has been made to compare the dissertations of Teachers College, Columbia University, with those of other institutions. The reader interested in making such comparisons is referred to the tables.

JUDGMENTS OF GRADUATE STUDENTS

During the year 1926-27, a course in Methods of Research was offered by several members of the faculty of Teachers College, Columbia University. At the close of the course, the members of the class were asked to vote on the topics that should be included in such a course the next year. All the group had had some experience in research, and some had completed their dissertations. The report of their judgments will be found in Table VII. Practically all the group, ninety-eight to be exact, participated. In voting, each student was influenced more or less by his own interests, and by the type of research in which he was engaged or which he proposed to undertake. The data in Table VII show that there was no unanimity of opinion as to the inclusion of any one phase of research. With the exception of "Selection of the Problem," each item in Table VII was checked by more than half of the group. A course in Methods of Research for a group with such varied needs and interests must clearly be comprehensive. This conclusion

TABLE VII

JUDGMENTS OF NINETY-EIGHT GRADUATE STUDENTS AS TO TOPICS THAT SHOULD BE INCLUDED IN A METHODS OF RESEARCH COURSE

	THOSE WHOSE DISSERTATIONS WERE			
	COM- PLETE	IN PROC- ESS	NOT YET BEGUN	TOTAL
NUMBER AND CLASSIFICATION OF JUDGES	6	23	69	98
Selection of the Problem	4	12	32	48
Preparation of Bibliography and Survey of Literature	4	10	37	51
Formulation of the Problem and Outline of Procedure	4	16	47	67
Examination of Proposed Sources of Data	5	16	49	70
Methods of Research				
Historical	4	15	35	54
Experimental	5	11	42	58
Techniques of Research				
Measurement				
Evaluation of existing tests	4	13	42	59
Construction of new tests	4	10	37	51
Questionnaires and interviews	4	14	42	60
Judgments or ratings	5	16	43	64
Statistical Treatment of Data				
Measures of central tendency	4	12	42	58
Measures of variability	4	10	41	55
Correlation methods	4	11	37	52
Measures of reliability	3	10	38	51
Presentation of a research	4	12	35	51

as to the content of a course in Methods of Research supports the conclusions based on the analysis of one hundred dissertations.

CHAPTER VIII

CONCLUSION

IN the preparation of this monograph, the author has had the following aims:

1. The construction of check lists for educational research, specifically quantitative research.
2. The application of these check lists to one hundred dissertations.

The items for the check lists have been obtained from books on the methodology of research, magazines, dissertations, and other sources. The various items have been arranged in five Check Lists, as follows:

1. Selection and Formulation of a Research Problem.
2. Methods of Research.
3. Techniques Used in Research.
4. Statistical Methods Used in Research.
5. The Presentation of the Research.

These Check Lists summarize, in convenient form, the procedures of research. Research workers may wish to insert additional items, notes, and bibliographical references. For this purpose, blank pages have been provided at the back of the book.

The Check Lists are intended as a practical and convenient aid or guide for at least two groups of persons:

1. Those who are inexperienced in research may use them in order not to overlook any essential procedures. The number of items found in the Check Lists is some indication of the complex nature of research.

2. Those who are directing research workers may use the Check Lists to enable them to keep in touch with the progress of many different research studies. The time required for interviews and directions should be considerably lessened.

In order to test the Check Lists, one hundred dissertations were read, and the procedures, techniques, and statistical methods used were systematically tabulated. These dissertations comprise seventy-three from Teachers College, Columbia University, and twenty-seven from other institutions. They constitute the dissertations of practically all the candidates in the field of education whose dissertations were printed during the years 1924 and 1925.

The Check Lists, as well as the tables showing the frequency with which the several techniques and procedures were used in actual dissertations, may also be used as a guide in training research workers. As an indication of relative emphasis, the following statements have been culled from the analyses of the one hundred dissertations:

1. In the statement of the problems a majority of the writers showed the worthwhileness of their problems, stated their objectives, and outlined the data to be collected.

2. About one-fifth of these research workers used historical methods of research; about one-third, experimental methods; and about three-fourths, investigational methods.

3. As a technique for collecting data, measurement was most frequently used. Thirty-eight per cent of the group used existing tests and twenty-three per cent constructed new tests. Questionnaires were used by forty-two per cent and ratings by twenty-three per cent

98 CHECK LISTS FOR EDUCATIONAL RESEARCH

of the group. A small number employed the interview technique, and a few, personal observation.

4. Statistical methods were used by practically every writer. Many complicated methods were found. More use could well have been made of reliability measures.

5. This group of writers used various methods of presenting their data, such as tables, graphs, charts, and the like. Summaries were frequently inserted to assist the reader.

The judgments of ninety-eight graduate students as to the topics that should be included in a course for training research workers were also obtained and are reported in tabular form.

All the evidence indicates that a course in Methods of Research should include instruction in practically all the available methods and techniques of research listed in the Check Lists.

BIBLIOGRAPHY ¹

1. ALEXANDER, CARTER. *Educational Research, Suggestions and Sources of Data with Specific Reference to Administration*. Bureau of Publications, Teachers College, Columbia University, 1927. 41 pp.
2. ALEXANDER, CARTER. "Opportunities for Research in Educational Finance." *Educational Administration and Supervision*, 9:209-22, April 1923. Also in the *Phi Delta Kappan*, 5:5-8, April 1923.
3. ALEXANDER, CARTER. "How to Equip the School Administrator with Methods of Locating Data Which Will Carry Over When He Is on the Job." *Teachers College Record*, 28:890-99, May 1927.
4. AYER, ADELAIDE. *Some Difficulties in Elementary School History*. Contributions to Education, No. 212. Bureau of Publications, Teachers College, Columbia University, 1926. 70 pp.
5. AYRES, LEONARD P. *An Index Number for State School Systems*. Russell Sage Foundation, New York, 1920. 78 pp.
6. BINGHAM, WALTER V. and FREYD, MAX. *Procedures in Employment Psychology*. A. W. Shaw Company, 1926. 262 pp.
7. BRINKLEY, STERLING G. *Values of New-Type Examinations in the High School, with Special Reference to History*. Contributions to Education, No. 161. Bureau of Publications, Teachers College, Columbia University, 1925. 122 pp.
8. BROOKS, FOWLER D. "Criteria of Educational Research." *School and Society*, 18:724-29, December 22, 1923.
9. BUCKINGHAM, BURDETTE R. *Research for Teachers*. Silver Burdette and Company, 1926. 386 pp.

¹ For the convenience of the research worker, this Bibliography has been limited to comparatively few references. Accessibility to the average research worker has been our aim. This has, necessarily, resulted in many arbitrary decisions and the elimination of many references that are equally good.

100 CHECK LISTS FOR EDUCATIONAL RESEARCH

10. BURGESS, W. R. "The Education of Teachers in Fourteen States." *Journal of Educational Research*, 3:161-72, March 1921.
11. BURGESS, W. R. "The Rate of Progress in Teacher Preparation." *Journal of Educational Research*, 4:180-86, October 1921.
12. BURKS, BARBARA S. "On the Inadequacy of the Partial and Multiple Correlation Technique." *Journal of Educational Psychology*, 17:532-40, November 1926 and 17:625-30, December 1926.
13. BURKS, BARBARA S. "Statistical Hazards in Nature-Nurture Investigations." *Twenty-seventh Yearbook, National Society for the Study of Education*, Part I:9-38. Public School Publishing Company, 1928. 465 pp.
14. BURTON, WM. H. "A Contribution to the Technique of Constructing 'Best-Answer' Tests." *Elementary School Journal*, 25:762-70, June 1925.
15. CARROLL, ROBERT P. "Methods of Educational Research." *Journal of Educational Research*, 11:309-11, April 1925.
16. CHADDOCK, R. E. *Principles and Methods of Statistics*. Houghton Mifflin Company, 1925. 471 pp.
17. CHAMBERS, GEORGE G. *An Introduction to Statistical Analysis*. F. S. Crofts and Company, 1925. 257 pp.
18. CHAPMAN, J. CROSBY. "Some Elementary Statistical Considerations in Educational Measurements." *Journal of Educational Research*, 4:212-20, October 1921.
19. CHAPIN, F. STUART. *Field Work and Social Research*. Century Social Science Series. The Century Company, 1920. 224 pp.
20. CHARTERS, W. W. "The Collecting of Unrecorded Specifics." *Journal of Educational Research*, 5:280-94, April 1922.
21. CHU, JENNINGS PINKWEI. *Chinese Students in America: Qualities Associated with Their Success*. Contributions to Education, No. 127. Bureau of Publications, Teachers College, Columbia University, 1922. 55 pp.
22. COURTIS, S. A. "The Relation between Rate and Quality in Educational Measurements." *Journal of Educational Research*, 10:110-31, September 1924.
23. COURTIS, S. A. *Why Children Succeed*. Courtis Standard Tests, Detroit, Michigan, 1925. 271 pp.

24. DEWEY, JOHN. *How We Think*. D. C. Heath and Company, 1910. 224 pp.
25. ENGELHARDT, N. L. *Superintendent's Checking List for the Evaluation of the Supervisory and Instructional Programs*. Preliminary mimeographed edition. Bureau of Publications, Teachers College, Columbia University, 1926. 33 pp.
26. ENGELHARDT, N. L., REEVES, C. E., and WOMRATH, G. F. *Score Card and Standards for Janitorial-Engineering Service*. Bureau of Publications, Teachers College, Columbia University, 1926. 53 pp.
27. FISHER, IRVING. *Purchasing Power of Money*. Macmillan Company, 1911. 505 pp.
28. FLING, FRED MORROW. *The Writing of History: An Introduction to Historical Method*. Yale University Press, 1923. 195 pp.
29. GARRETT, HENRY E. *Statistics in Psychology and Education*. Longmans, Green and Company, 1926. 317 pp.
30. GARRETT, HENRY E. "An Empirical Study of the Various Methods of Combining Incomplete Order of Merit Ratings." *Journal of Educational Psychology*, 15:157-71, March 1924.
31. GATES, ARTHUR I. "The Relation of Quality and Speed of Performance: A Formula for Combining the Two in the Case of Handwriting." *Journal of Educational Psychology*, 15:129-44, March 1924.
32. GATES, ARTHUR I. "A Modern Systematic versus an Opportunistic Method of Teaching." *Teachers College Record*, 27:679-700, April 1926.
33. GRAY, WM. S. "Problems for Scientific Investigation in Elementary Education." *Elementary School Journal*, 23: 654-63, May 1923.
34. HAMILTON, FREDERIC RUTHERFORD. *Fiscal Support of State Teachers Colleges*. Contributions to Education, No. 165. Bureau of Publications, Teachers College, Columbia University, 1924. 51 pp.
35. HARPER, ERNEST B. "The Case Study Method." *Religious Education*, 22:505-12, May 1927.
36. HOLZINGER, KARL. "An Analysis of Errors in Mental Measurement." *Journal of Educational Psychology*, 14:278-88, May 1923.
37. HOLZINGER, KARL. "Formulas for the Correlation between

102 CHECK LISTS FOR EDUCATIONAL RESEARCH

- Ratios." *Journal of Educational Psychology*, 14:344-47, September 1923.
38. HOLZINGER, KARL. "Scoring Multiple Response Tests." *Journal of Educational Psychology*, 15:445-47, October 1924.
39. HOLZINGER, KARL AND CLAYTON, BLYTHE. "Further Experiments in the Application of Spearman's Prophecy Formula." *Journal of Educational Psychology*, 16:289-99, May 1925.
40. KELLEY, TRUMAN L. *Statistical Method*. Macmillan Company, 1924. 390 pp.
41. KELLEY, TRUMAN L. *Interpretation of Educational Measurements*. World Book Company, 1927. 363 pp.
42. KULP, DANIEL H. *Outlines of the Sociology of Human Behavior*. A. G. Seiler (New York City). 1926. 257 pp.
43. LEE, BALDWIN. "Some Faults Common in Informal Objective Tests Made by High School Teachers." *Educational Administration and Supervision*, 14:105-13, February 1928.
44. MATHEWS, C. O. "The Effect of Position of Printed Response Words upon Children's Answers to Questions in Two-Response Types of Tests." *Journal of Educational Psychology*, 18:445-57, October 1927.
45. MCCALL, WM. A. *How to Measure in Education*. Macmillan Company, 1922. 416 pp.
46. MCCALL, WM. A. *How to Experiment in Education*. Macmillan Company, 1923. 281 pp.
47. MCCALL, WM. A. "Construction of the Multi-Mental Scale." *Teachers College Record*, 27:394-415, January 1926.
48. MCCALL, WM. A. AND BIXLER, HAROLD H. *How to Classify Pupils*. Bureau of Publications, Teachers College, Columbia University, 1928. 83 pp.
49. MCGAUGHY, JAMES RALPH. *The Fiscal Administration of City School Systems*. Macmillan Company, 1924. 95 pp.
50. MILLS, FREDERICK C. *Statistical Methods Applied to Economics and Business*. Henry Holt and Company, 1924. 604 pp.
51. MONROE, WALTER S. *Theory of Educational Measurements*. Houghton Mifflin Company, 1923. 364 pp.
52. MONROE, WALTER S. AND ASHER, OLLIE, *A Bibliography*

- of Bibliographies.* University of Illinois Bulletin, Vol. 24, No. 44, Bureau of Educational Research Bulletin No. 36. University of Illinois, Urbana, 1927. 60 pp.
53. MONROE, WALTER S. AND ENGELHART, MAX D. *The Techniques of Educational Research.* University of Illinois Bulletin, Vol. 25, No. 19, Bureau of Educational Research Bulletin No. 38. University of Illinois, Urbana, 1928. 84 pp.
54. MOORE, CLYDE B. *Civic Education.* Contributions to Education, No. 151. Bureau of Publications, Teachers College, Columbia University, 1924. 119 pp.
55. ODELL, C. W. *Educational Statistics.* The Century Company, 1925. 334 pp.
56. O'REAR, F. B. *The Duties of the Registrar.* Bulletin, Vol. 21, No. 1. Southwest Missouri State Teachers College, Springfield, Mo., January 1926. 172 pp.
57. OTIS, A. S. *Statistical Method in Educational Measurement.* World Book Company, 1925. 332 pp.
58. PEARSON, KARL. *The Grammar of Science.* Macmillan Company, 1911. 394 pp.
59. PRESSEY, S. L. *Mental Abnormality and Deficiency.* Macmillan Company, 1926. 356 pp.
60. REEDER, WARD G. *How to Write a Thesis.* Public School Publishing Company, 1925. 132 pp.
61. *Register of Teachers College Doctors of Philosophy 1899-1925.* Teachers College, Columbia University, 1925. 40 pp.
62. *Report of the Joint Committee on Standards for Graphic Presentation.* American Society of Mechanical Engineers, New York, 1918. 50 pp.
63. "Research Problems Proposed." *The Phi Delta Kappan*, 9:33-53, 76-78, 88, October and December 1926.
64. RICHMOND, MARY E. *Social Diagnosis.* Russell Sage Foundation, New York, 1917. 511 pp.
65. RUCH, G. M. "Minimum Essentials in Reporting Data on Standard Tests." *Journal of Educational Research*, 12: 349-58, December 1925.
66. RUCH, G. M. AND DEGRAFF, MARK H. "Corrections for Chance and 'Guess' vs. 'Do not guess' Instructions in Multiple-Response Tests." *Journal of Educational Psychology*, 17:368-75, September 1926.
67. RUCH, G. M. AND STODDARD, GEORGE D. *Tests and Meas-*

104 CHECK LISTS FOR EDUCATIONAL RESEARCH

- urements in High School Instruction.* World Book Company, 1927. 381 pp.
68. RUGG, H. O. *Statistical Methods Applied to Education.* Houghton Mifflin Company, 1917. 410 pp.
 69. RUGG, H. O. "Is the Rating of Human Character Practicable?" *Journal of Educational Psychology*, 12:425-38, November 1921; 485-501, December 1921; 13:30-42, January 1922; 81-93, February 1922.
 70. SCHLUTER, WM. C. *How to Do Research Work.* Prentice-Hall, Inc., 1926. 134 pp.
 71. SHEFFIELD, ADA E. *The Social Case History.* Russell Sage Foundation, New York, 1920. 227 pp.
 72. SHEN, EUGENE. "The Standard Error of Certain Estimated Coefficients of Correlation." *Journal of Educational Psychology*, 15:462-65, October 1924.
 73. SHEN, EUGENE. "A Note on the Standard Error of the Spearman-Brown Formula." *Journal of Educational Psychology*, 17:93-4, February 1926.
 74. SNEDDEN, DAVID. *Educational Sociology.* The Century Company, 1924. 689 pp.
 75. STRAYER, G. D. "The Scientific Approach to the Problems of Educational Administration." *School and Society*, 24: 685-95, December 4, 1926.
 76. STRAYER, G. D. AND ENGELHARDT, N. L. *Score Card and Standards for Elementary School Buildings.* Bureau of Publications, Teachers College, Columbia University, 1923. 64 pp.
 77. STRAYER, G. D. AND ENGELHARDT, N. L. *Score Card and Standards for High School Buildings.* Bureau of Publications, Teachers College, Columbia University, 1924. 95 pp.
 78. STRAYER, G. D., ENGELHARDT, N. L., AND ELSBREE, W. S. *Score Card and Standards for the Administration Building of a School System.* Bureau of Publications, Teachers College, Columbia University, 1927. 42 pp.
 79. STRAYER, G. D. AND OTHERS. *Report of the Survey of the Schools of Beaumont, Texas.* Bureau of Publications, Teachers College, Columbia University, 1927. 337 pp.
 80. STRAYER, G. D. AND OTHERS. *Report of the Survey of the Schools of Jacksonville, Florida.* Bureau of Publications, Teachers College, Columbia University, 1927. 433 pp.

81. STRAYER, G. D. AND OTHERS. *Report of the Survey of the Schools of Lynn, Massachusetts*. Bureau of Publications, Teachers College, Columbia University, 1927. 368 pp.
82. SYMONDS, P. M. "On the Loss of Reliability in Ratings Due to Coarseness of the Scale." *Journal of Experimental Psychology*, 7:456-61, December 1924.
83. SYMONDS, P. M. "Variations of the Product-Moment (Pearson) Coefficient of Correlation." *Journal of Educational Psychology*, 17:458-69, October 1926.
84. SYMONDS, P. M. *Measurement in Secondary Education*. Macmillan Company, 1927. 588 pp.
85. THORNDIKE, EDWARD L. *Mental and Social Measurements*. Bureau of Publications, Teachers College, Columbia University, 1913. 271 pp.
86. THORNDIKE, EDWARD L. "The Gains Made in Ability in English by Pupils Who Study Latin and by Pupils Who Do Not." *School and Society*, 18:690, December 8, 1923.
87. THORNDIKE, EDWARD L. *The Measurement of Intelligence*. Bureau of Publications, Teachers College, Columbia University, 1926. 616 pp.
88. TOOPS, H. A. "Validating the Questionnaire Method." *Journal of Personnel Research*, 2:153-69, 1923.
89. TOUTON, FRANK C. "Formulas for the Determination of Certain Measures of Central Tendency and Variability for a Composite Group." *Journal of Educational Psychology*, 16:170-74, March 1925.
90. TRABUE, M. R. "Educational Research in 1925." *Journal of Educational Research*, 13:336-44, May 1926.
91. UNITED STATES LIBRARY OF CONGRESS. *A List of American Doctoral Dissertations, Printed in 1924 and 1925*. Superintendent of Documents, Washington, D. C., 1927. 206 pp.
92. WATSON, GOODWIN B. "A Supplementary Review of Measures of Personality Traits." *Journal of Educational Psychology*, 18:73-87, February 1927.
93. WEIDEMANN, C. C. *How to Construct the True-False Examination*. Contributions to Education, No. 225. Bureau of Publications, Teachers College, Columbia University, 1926. 118 pp.
94. WHIPPLE, GUY M. "The Improvement of Educational Research." *School and Society*, 26:249-59, August 27, 1927.

106 CHECK LISTS FOR EDUCATIONAL RESEARCH

95. WILLIAMS, J. HAROLD. *Graphic Methods in Education*. Houghton Mifflin Company, 1924. 319 pp.
96. WINDES, E. E. and GREENLEAF, W. J. *Bibliography of Secondary Education Research 1920-1925*. Bureau of Education Bulletin, 1926, No. 2. Superintendent of Documents, Washington, D. C. 95 pp.
97. WINDES, E. E. *Bibliography of Studies in Secondary Education*. Bureau of Education Bulletin, 1927, No. 27. Superintendent of Documents, Washington, D. C. 30 pp.
98. WOOD, BEN D. "Studies of Achievement Tests." *Journal of Educational Psychology*, 17:1-22, January 1926, and 17:125-39, February 1926.
99. WYLIE, ANDREW T. "To What Extent May We Rely upon the Answers to a School Questionnaire?" *Journal of Educational Method*, 6:252-57, February 1927.
100. YULE, G. UDNY. *An Introduction to the Theory of Statistics*. C. Griffin and Company, London (Fifth Edition, enlarged), 1919. 382 pp.

APPENDIX

ADDITIONAL ITEMS, NOTES, AND REFERENCES

108 CHECK LISTS FOR EDUCATIONAL RESEARCH

Additional Items, Notes, and References to be recorded here.

Additional Items, Notes, and References to be recorded here.

110 CHECK LISTS FOR EDUCATIONAL RESEARCH

Additional Items, Notes, and References to be recorded here.

Additional Items, Notes, and References to be recorded here.

112 CHECK LISTS FOR EDUCATIONAL RESEARCH

Additional Items, Notes, and References to be recorded here.

INDEX

INDEX

- Administration of a test, 35
Age norms, 31
Analysis, of classroom procedure, 50; of court decisions, 50; of data, 52; of minutes, 50; of pupil performance, 50; of school records, 50; of statutes, 50; of study habits, 50; of textbooks, 50
Appendix, 76
Application of the check lists, 3, 80
Assumptions, of a normal distribution, 53; of random sampling, 17; underlying statistical formulas for computing measures of reliability, 58; underlying statistical procedures, 53
Attenuation, correction for, 56
Average deviation, 55, 90

Bibliographies, available, 6
Bibliography, 75, 93; preparation of, 6; sources to be consulted in preparing, 7

Calibration of a test, 31
Case method, 21
Case-group method, 22
Causal investigation, 23
Central tendency, measures of, 53, 90; measures of, for a composite group, 55
Checking the data, 50
Choice of material for tests, criteria for, 26
Citations, 71

Class intervals, number of, 47, 53
Coarseness of grouping data, correction for, 56
Coefficient of correlation, 56; interpretation of, 56; spurious, 56, 57
Coefficient of reliability, 59
Conclusion, 73, 96
Constant errors, effect of not measured by reliability formulas, 60; in rating, 45
Construction of the check lists, 2
Correction, for attenuation, 56; for coarseness of grouping, 56
Correlation, bi-serial r method, 56, 90; contingency method, 56, 90; correlation ratio (η) method, 56, 90; methods of computing, 56; multiple correlation method, 56, 57, 90; partial correlation method, 56, 57, 90; of chance halves of a test, 60; of scores obtained by repetition of the same form of a test, 60; of similar forms of a test, 60; product-moment method, 56, 90; rank, 56, 90; tables, 56

Data, presentation of, 65; sources of, 11, 14; statistical treatment of, 52
Decile, 54, 90
Definition of the problem, 64
Descriptive investigation, 23

- Deviation, average, 55, 90; median, 90; quartile, 54, 90; standard, 55, 90
- Diagnostic value of a test, 34
- Dissertations, analysis of, 81
- Distribution, frequency, 53; normal, 53; order, 53; rank, 53
- Duplicate forms of a test, 33
- English structure and form, 78
- Equating groups, 19
- Equivalent - groups experiment, 16, 85
- Errors, probable, 58, 91; response, 61; sampling, 61; scale, 61; scoring, 61
- Experiment, equivalent-groups, 16, 85; materials for, 18; one-group, 16, 85; place of, 18; rotation, 16, 85; subjects of an, 16; time of, 19
- Experimental factors, control and measurement of, 19
- Experimental methods, 15, 85
- Formulation of a research problem, 9
- Grade norms, 31
- Grade scores, 31
- Graduate students, judgments of, 94
- Graphic rating scale, 46
- Graphs, 69, 93
- Groups, equating, 19
- Halo effect, 45, 49
- Historical methods of research, 13, 85
- Incomplete order of merit rankings, 62
- Index numbers, 62; reliability, 59, 91
- Indices, correlation of, 57
- Interviewers, 39
- Interviews, 38, 87
- Introduction for a research, 63
- Investigation, causal, 23; descriptive, 23
- Investigational methods, 23
- Irrelevant factors, control and measurement of, in experiments, 20
- Judges, 48
- Judgments, 45, 87
- Log, 21, 51
- Man - to - man comparison scale, 46
- Manual of directions for a test, 32
- Materials for an experiment, 18
- Mean, 54, 90
- Measurement, 25, 87
- Measures, of central tendency, 53, 90; of central tendency for a composite group, 55; of relationship, 55; of variability, 54, 90; of variability for a composite group, 55
- Mechanical features of a test, 34
- Median, 54, 90
- Methods of research, 13, 85; case, 21, 85; case-group, 22, 85; experimental, 15, 85; historical, 13, 85; in one hundred dissertations, 85; investigational methods, 23, 85
- Mode, 54, 90

- Norms, age, 31; approximate, 31; grade, 31; percentile, 31; sigma position, 31; test, 31
- Observation, 50, 87
- One-group experiment, 16, 85
- Order of merit ranking, 46
- Outline of procedure, 9
- Paired comparisons, 46
- Percentile, 54, 90
- Percentile norms, 31
- Place of the experiment, 18
- Presentation, form of, 76, 93; method of, in one hundred dissertations, 93; of a research, 63, 93; of data, 65
- Probable error, 58, 91; of a coefficient of correlation (r), 59, 91; of a difference, 59, 91; of a mean, 58, 91; of a median, 58, 91; of a percentile, 58; of a score, 58, 90; of a standard deviation, 58; of estimate, 59, 91; of percentage measures of overlapping, 59; of Q , 58
- Problem, definition of, 64; formulation of, 9; methods of approach of, 10; statement of, 64, 83; survey of literature concerning, 6, 83
- Procedure, outline of, 9
- Qualitative analysis of data, 52
- Quartile, 54, 90
- Quartile deviation, 54, 90
- Questionnaires, 40, 87; arrangement of items, 41; follow-up, 44; form, 43; mailing, 44; selection of items, 41
- Range, 54
- Ranking, 48
- Raters, 48
- Ratings, 45, 87; methods of obtaining, 46; number required for reliability, 60; on personality traits, 46; reliability of, 46, 48; scales for, 46; self-ratings, 47; transmutation into quantitative units, 46; validity of, 48
- References, 72, 93
- Regression equations, 58
- Reliability, coefficients of, 28, 59; index of, 59, 91; measures of, 58, 91; of a test, 28, 60; of a test, facts to be reported concerning, 60
- Research, introduction for, 63; methods of, 13, 85; methods of, in one hundred dissertations, 85; presentation of, 63, 93
- Rotation experiment, 16, 85
- Scales, number of class intervals, 47; rating, 46
- Scaling a test, 30
- Score card, 50
- Scoring, formulas and keys, 33; of tests, 37
- Selection of a problem, 5
- Selection of the researches for analysis, 81
- Self-ratings, 47
- Sheppard's correction to the standard deviation, 55
- Sigma, 55; position, 31
- Sources, evaluation of, 14; for American educational history, 14
- Sources of data, primary, 11, 14; secondary, 11
- Spearman - Brown formula, 29, 47, 59, 60

- Standard deviation, 55, 90
Standard measures, 62
Statement of the problem, 64, 83; in one hundred dissertations, 84
Statistical computations in one hundred dissertations, 89
Statistical methods, 52
Subjects of an experiment, 16
Summaries, 73, 93; of previous or related studies, 9, 65
Survey, of literature concerning problem, 6; school, 50
Tables, 67, 93
Techniques, of collecting data in one hundred dissertations, 87; used in research, 25
Tests, administration of, 35; diagnostic value of, 34; duplicate forms, 33; keys, 33; manual of directions for, 32; mechanical features, 34; norms, 31; reliability of, 28, 60; scaling, 30; scoring, 37; scoring formulas for, 33; validity, 26
Time of an experiment, 19
Transmutation of ratings into quantitative units, 46
T scores, 61, 91
Use of the check lists, 4
Validity of a test, 26
Variability, measures of, 54; measures of, for a composite group, 55; measures of relative, 55
Variation, coefficient of, 55, 90

University of St. Francis
GEN 378 B576
Bixler
Check Lists for Educational Re



3 0301 00018345 5

07-DFI-895

